

特別研究論文

(査読済み)

研 究 題 目

音楽信号処理における基本周波数推定を応用した心拍信号解析

提出年月日	2022 年	1 月	31 日
氏 名	梶谷 奈未		
主 査	北村 大地 講師		
副 査	重田 和弘 教授		
副 査	石井 耕平 准教授		

香川高等専門学校
専攻科
創造工学専攻



Heart Rate Analysis Exploiting Fundamental Frequency Estimation in Music Signal Processing

Nami Kajitani

Advanced Course in Industrial and Systems Engineering
National Institute of Technology, Kagawa College

Abstract

In home health care, it is important to remotely monitor patients and to manage their health condition. For this purpose, patients need to wear sensors to acquire information such as heart rate and blood pressure, which is cumbersome and leads to a decrease in the quality of life. Therefore, a nail tip sensor has been developed as a wearable sensor that can be securely fixed for long periods of time without discomfort. Since this sensor can obtain photoplethysmogram (PPG) from changes in blood flow at the fingertip, it is expected to be applied to the constant monitoring of heart rate in real environments. However, since the signal obtained from this sensor contains a large amount of noise caused by body movements, it is necessary to extract only the heart rate component from the waveform by signal processing. Little research has been conducted to determine the best method for heart rate estimation with PPG obtained from the nail tip sensor. In music signal processing, the problem of estimating pitch (fundamental frequency) from observed signals such as vocal and instrumental sounds is a major topic, and various methods have been proposed so far. The estimation problem of the fundamental frequency itself is often encountered not only in acoustic signals but also in many other media signals. In biological signal processing, heart rate estimation is essentially equivalent to an estimation of fundamental frequency. Thus, there is a possibility that the methods in music signal processing can be applied to the heart rate estimation. In this thesis, in order to estimate the heart rate from PPG obtained by the nail tip sensor, I applied four methods, the normalized autocorrelation function method, the cepstrum method, the YIN method, and the multiple signal classification method, which are typical fundamental frequency estimation methods in music signal processing. The effectiveness of these methods was experimentally verified. As a result, it was found that the fundamental frequency estimation method in music signal analysis is useful to some extent for the analysis of PPG.

Key Words: fundamental frequency estimation, music signal analysis, health care

(和訳)

在宅医療では、患者の健康状態を管理するために、遠隔でモニタリングを行うことが重要である。この目的のために、患者は心拍や血圧等の情報を取得するセンサを装着し続ける必要があり、その煩雑さから生活の質の低下を招く問題がある。そこで、不快な装着感が無く長期間にわたる強固な固定が可能なウェアラブルセンサとして、付け爪型センサが提案されている。このセンサでは、指先の血流の変化によって光電脈波信号 (PPG) が得られるため、実生活環境下での常時心拍モニタリングへの応用が期待されている。しかしながら、このセンサから得られる信号には体動に起因するノイズが多く含まれているため、信号処理を施すことで波形から心拍成分のみを抽出する必要があるが、付け爪型センサで得られる PPG に対してどのような手法が心拍推定に最良であるかについては、これまであまり調査が進められていない。一方で、音楽信号解析において、ボーカルや楽器音等の観測信号から音の高さ (基本周波数) を推定する問題は一大トピックであり、これまで多様な手法が提案されている。本来、信号の基本周波数の推定問題自体は、音響信号のみならず多くのメディア信号で見られる問題であり、生体信号処理においては計測波形から心拍を推定する問題も本質的に基本周波数推定問題である。そのため、基本周波数推定問題と等価である心拍推定に対して、音楽信号解析の手法が応用できる可能性がある。そこで、本研究では、付け爪型センサで得られる PPG から心拍を推定するために、音楽信号解析において代表的な基本周波数推定法である正規化自己相関関数法、ケプストラム法、YIN 法、及び多重信号分類法の 4 つの手法を適用し、生体信号の解析における有効性を検証した。その結果、音楽信号解析における基本周波数推定法は生体信号の解析に、ある程度有用であることが分かった。

目次

第 1 章	緒言	1
1.1	本論文の背景	1
1.2	本論文の目的	3
1.3	本論文の構成	3
第 2 章	付け爪型センサの概要	6
2.1	まえがき	6
2.2	特徴	6
2.3	動作原理	7
2.4	測定波形	7
2.5	本章のまとめ	10
第 3 章	音楽信号解析における代表的な基本周波数推定法	13
3.1	まえがき	13
3.2	本研究の動機	13
3.3	信号の前処理	14
3.4	正規化自己相関関数法に基づく心拍推定	14
3.5	ケプストラム法に基づく心拍推定	16
3.6	YIN 法に基づく心拍推定	18
3.7	MUSIC 法に基づく心拍推定	18
3.8	本章のまとめ	22
第 4 章	実験	23
4.1	まえがき	23
4.2	評価指標	23
4.3	データ 1 に対する実験結果	24
4.3.1	実験条件	24
4.3.2	実験結果比較	26
4.4	データ 2 に対する実験結果	26
4.4.1	実験条件	26
4.4.2	実験結果比較	26

4.5	データ 3 に対する実験結果	27
4.5.1	実験条件	27
4.5.2	実験結果比較	28
4.6	本章のまとめ	29
第 5 章	結言	33
	謝辞	34
	参考文献	34

第 1 章

緒言

1.1 本論文の背景

高齢社会となった日本では、在宅医療の重要性が年々増加している。在宅医療では、患者の健康管理のために遠隔でモニタリングを行うことが重要であるが、この実現のためには、Fig. 1.1 に示すように、センサを身体に装着することで心拍や血圧等を取得する必要がある。現在では、Fig. 1.2 のような腕時計式や胸部貼り付け式などのウェアラブルセンサが販売されており、これらを装着することで遠隔モニタリングが行われている。しかし、これらの機器には、不快な装着感や入浴時の着脱の必要性、皮膚表面の蒸れに起因する衛生面の問題、及び接触皮膚炎の誘発などの問題が生じる。

そこで現在、上記問題への対策として、付け爪型センサの開発が進められている [1]。これは Fig. 1.3 のように、ジェルネイル等を用いて爪の表面にセンサを固定する小型デバイスであり、脈拍や心拍の計測を目的としている。爪のサイズの回路にセンサ、バッテリー、送信機等を搭載しており、指先の血管の拡張・収縮に起因する爪表面の色変化を捉えた光電脈波信号 (photoplethysmogram: PPG) が計測できる。この PPG を無線で送信することで、心拍を遠隔で常時記録・解析するシステムを想定している (Fig. 1.3 では有線で接続しているが、本来は無線で送信するため、爪の上だけで完結するデバイスとなる)。このセンサシステムの概要を Fig. 1.4 に示す。文献 [1] で提案されている付け爪型センサの利点は以下の 3 点である。

- 爪には感覚神経が無いため、不快な装着感がない
- 爪は組織が固いため、数週間固定を維持できる
- 複数の爪に装着することで多点同時計測ができ、より発展的な計測が期待できる

前述のように、多くの利点を持つ付け爪型センサはその実用性の高さから実社会への応用が期待されている。しかしながら、爪から観測される PPG には、体動等に由来するノイズが観測されてしまう問題があるため、ノイズに対して頑健な心拍推定法が必要となる。また、このような PPG に対して効果的な心拍推定法はあまり調査されておらず、適切なアルゴリズムは明らかとなっていないのが現状である。

一方で、音響信号処理の領域で活発に研究されている音楽信号解析という分野では、ボーカ

2 第1章 緒言

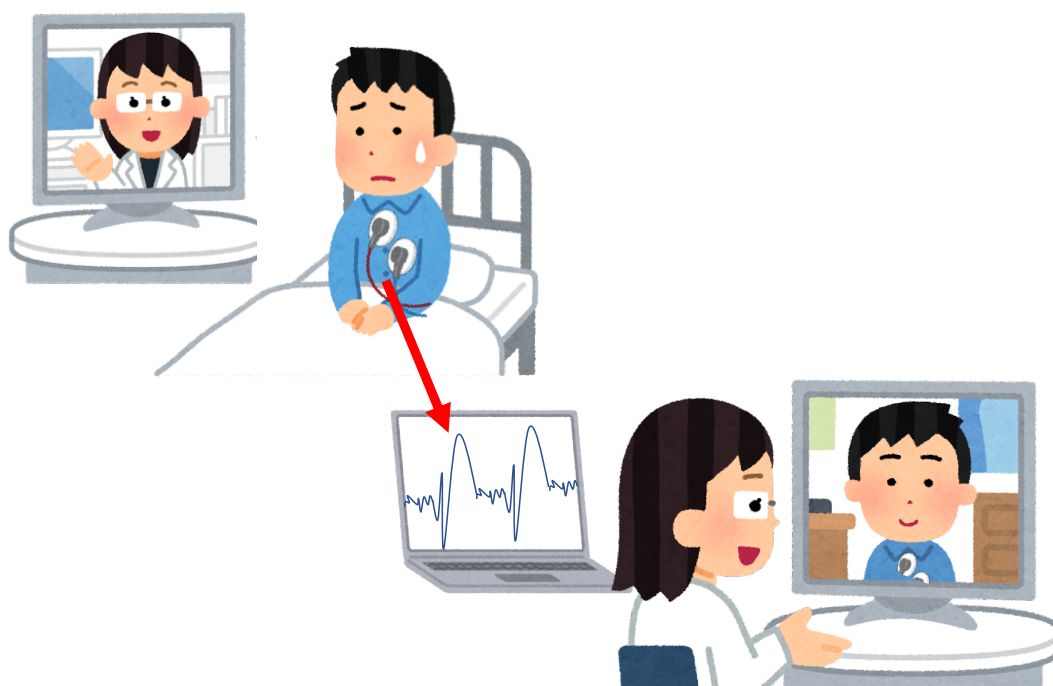


Fig. 1.1. Telemedicine system using contact type heart rate monitor system.

ルや楽器音等の音響信号から音の高さ（ピッチ又は基本周波数と呼ばれる）を推定する問題が一大トピックとなっている。これは、音の基本周波数を推定することで、メロディの追跡やコード推定等の技術に応用することができるためであり、基本周波数は音響信号を特徴づけるきわめて重要な要素であることに起因している。そのため、音楽信号解析分野では、これまで多様な基本周波数推定法が提案されている。例えば、時間波形の自己相関関数を算出して信号の基本周期を求める正規化自己相関関数法 [2], 信号の対数振幅スペクトルの周波数成分を解析するケプストラム法 [3], 話者やボーカル等, 人間の音声の基本周波数を頑健に推定することができる YIN 法^{*1}[4], 及び部分空間法に基づく信号の分類手法である多重信号解析 (multiple signal classification method: MUSIC) 法 [5] などがある。これらの基本周波数推定法は、前述の音楽信号解析の前処理として有用性が認められている。本来、信号の基本周波数の推定問題自体は、音響信号のみならず多くのメディア信号で見られる問題であり、生体信号処理においては計測波形から心拍を推定する問題も本質的に基本周波数推定問題である。そこで本研究では、付け爪型センサで得られる PPG から心拍を推定するために、音楽信号解析において代表的な、前述の 4 つの基本周波数推定法を適用し、生体信号の解析における有効性を検証する。

^{*1} YIN 法の YIN は東洋の陰と陽に由来する語句であり、何らかの英単語の省略形ではない。



Fig. 1.2. Wearable devices used for remote monitoring [6, 7].

1.2 本論文の目的

本節では、本研究の目的について説明する．医療工学の分野では、接触型心拍センサ等で測定される心電信号（electrocardiograph: ECG）に対して、波形の基本周期（基本周波数の逆数）を算出することで心拍を推定することが一般的である．この波形の基本周期は、ECGに含まれる R 波と呼ばれるスパイク状の波形（Fig. 1.5 の ECG 波形を参照）の時間間隔を推定する手法が良く用いられ、R-R 間隔（R-R interval: RRI）推定と呼ばれる．付け爪型センサが提案された文献 [1] では、RRI 推定の代替として、ゼロクロス法と呼ばれる簡易的なアルゴリズムを用いて、付け爪型センサから得られた PPG 及び接触型心拍センサで得られた ECG の基本周期を推定している．ゼロクロス法の概要を Fig. 1.5 に示す．ゼロクロス法とは、正負が反転しているサンプルの間隔を信号の基本周期として求める手法であり、正負反転を検知しないデッドタイムを手動で設定する必要がある．本論文では、付け爪型センサで測定される PPG に対して、よりノイズに頑健な心拍推定法を目指し、ゼロクロス法以外の心拍推定法を適用する．特に、前節で述べた音楽信号解析分野における代表的な 4 手法を適用し、推定精度について議論する．

1.3 本論文の構成

本節にて、本論文の構成を示す．まず、2 章では、本論文で取り扱う、心拍を測定する付け爪型センサと呼ばれる装置の概要について詳細に解説する．付け爪型センサを利用することで、これまでのウェアラブルセンサにおいて問題視されていた不快な装着感を改善でき、多点同時計測の実現や、ジェルネイルを用いることで長期間にわたる強固な固定も可能となる．3

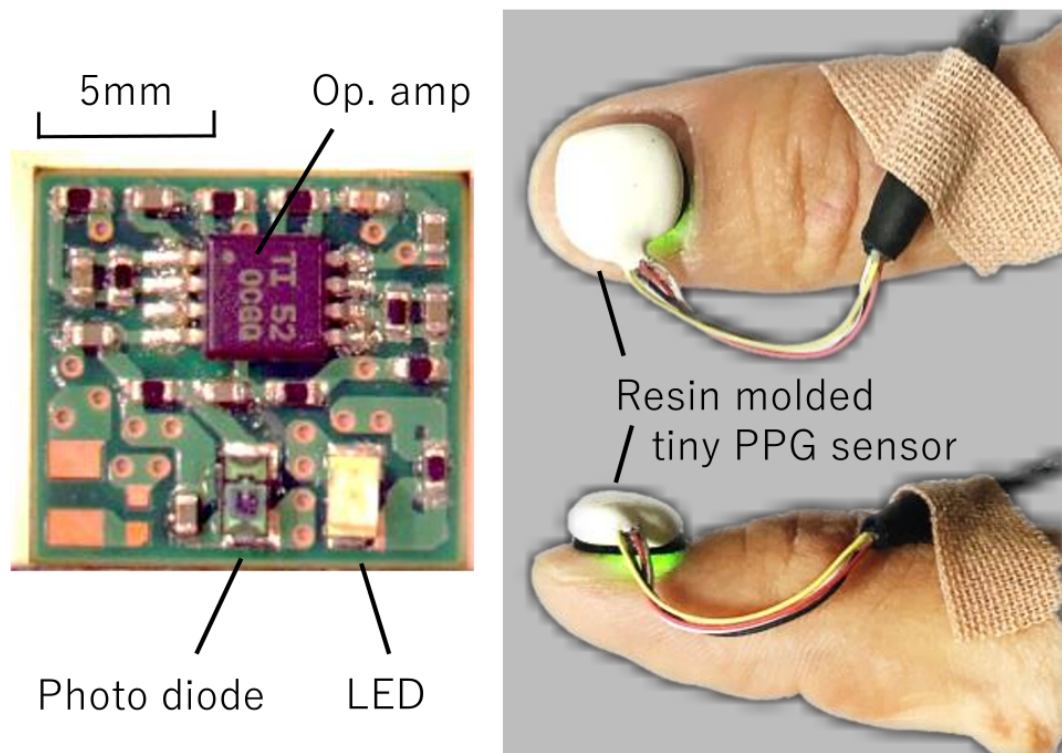


Fig. 1.3. Nail tip sensor for PPG measurement [1].

章では、本論文で実際に活用する4つの基本周波数推定法の推定法について詳細に解説する。いずれの手法も音楽信号処理の分野では基本周波数推定法として一般的に用いられる。4章では、PPGに対して音楽信号解析で用いられる4手法を適用した際の推定精度の比較を行う。複数の観測データを用いて、それぞれの手法の精度及び特徴を統計的な観点から比較する。最後に、5章では、本論文の結論を述べる。

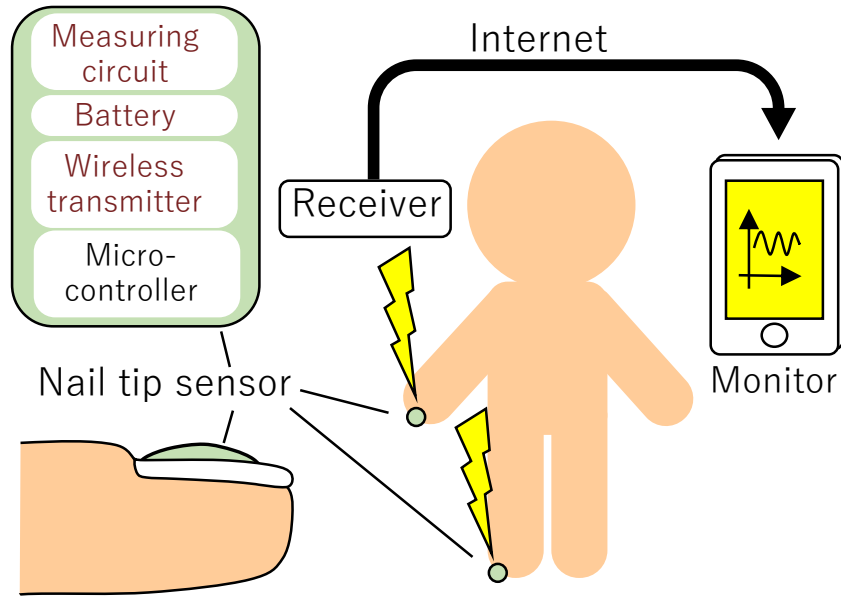


Fig. 1.4. Nail tip sensor system [1].

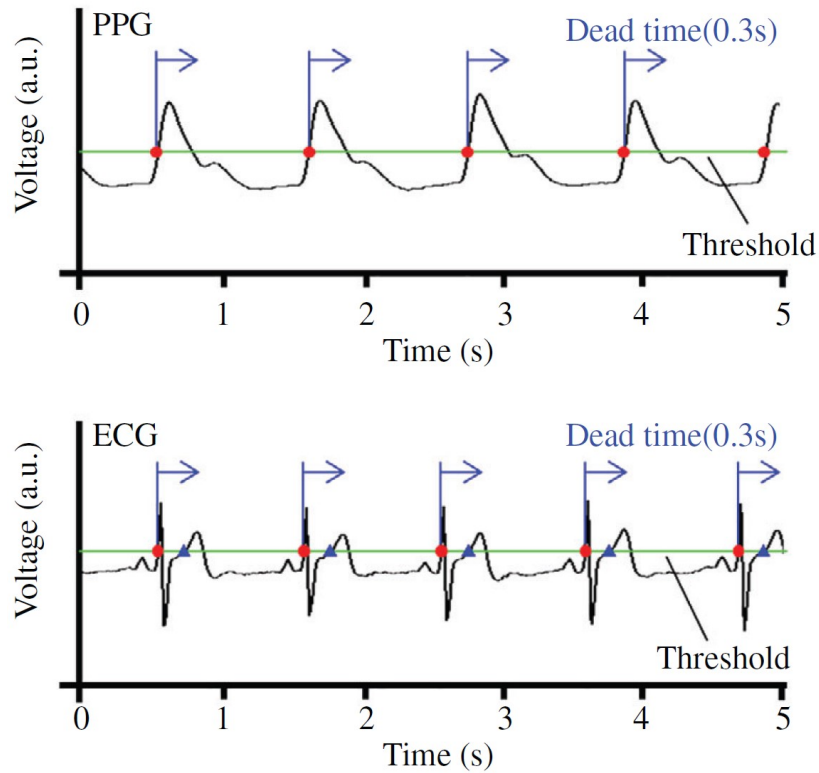


Fig. 1.5. Overview of zero-cross method. Time interval between zero-crossed samples is calculated, where dead time is utilized for avoiding miss estimation.

第 2 章

付け爪型センサの概要

2.1 まえがき

本章では、本研究で用いる付け爪型センサ [1] の概要を説明する。まず 2.2 節では、付け爪型センサの特徴について詳細に述べる。ここでは、付け爪型センサの開発の重要性や、本論文で問題とされている外乱の特徴についても説明する。2.3 節では、付け爪型センサの計測原理について述べる。計測原理では、問題視されている外乱の発生原理についても詳しく解説する。2.4 節では、付け爪型センサから得られた測定波形について述べる。この測定波形が本研究で取り扱う入力信号となる。最後に、2.5 節では、本章のまとめを述べる。

2.2 特徴

在宅医療において、患者の心拍をリアルタイムで遠隔モニタリングすることは必須である。しかし、心拍測定を目的として一般的に用いられるセンサは、身体に取り付ける電極によって、不快な装着感や着脱のストレス、及び衛生面等の多くの問題が生じる。一方で、文献 [1] で提案されている付け爪型センサでは、爪には感覚神経が無いことや、組織が固く長期間に渡る固定が維持できることから、一般的に用いられているセンサに生じる問題を全面的に解消できる。なお、本センサでは、センサの固定のためにジェルネイルを用いているため、数週間程度の固定も実現できる。さらに、人間の爪は両手両足で合計 20 枚存在するため、将来的には複数の爪に装着することで多点同時計測も可能となり、より正確にモニタリングすることができると考えられる。これが実現できれば、より発展的な計測も期待でき、心拍測定のみでなく、血圧測定にも応用できると予想される。

血液は心拍に合わせて拡張や収縮を周期的に繰り返しながら血管を流れている。そこで、本センサでは、爪が皮膚に接触している面（爪床）にある毛細血管の拡張及び収縮から生じる微小な色の変化を調べることで心拍を推定している。ところが、爪周辺の皮膚は非常に繊細であるため、体動や筋動に起因する外乱からも、色の変化が同時に生じている。なお、血管の拡張及び収縮に起因する色の変化は非常に微弱であるため、爪から得られる測定波形は、前述の外

乱によって心拍成分がほとんど観測できない状態になる。本研究では、この問題を軽減するために、爪から得られた PPG に対して、音楽信号解析分野で発展した基本周波数推定法を適用し、ノイズに頑健な心拍推定が行えるかについて詳しく調査する。本論文では、以後 PPG 中の心拍に由来する波形成分を心拍信号と呼び、体動や筋電等の心拍以外に由来する波形成分をノイズと呼ぶ。

2.3 動作原理

本研究で用いる付け爪型センサの動作原理は Fig. 1.4 に示されている通りである。付け爪型センサ本体には脈波計測回路、バッテリー、無線送信機、及び制御用マイコンが搭載されている。なお、ここで記載されている無線送信機について、Fig. 1.3 では有線で接続しているが、将来的には無線通信で送信できるように改良予定である。

2.2 節でも説明したように、爪先では血管の拡張及び収縮によって微細な色の変化が生じる。本センサでは、これによって変動した光の反射率を利用することで脈波を計測している。爪の表面から発せられた光は爪を透過し、爪床に到達する。ここで、爪床の反射光をフォトダイオード検出する。この反射光は、増幅回路を通過した後に出力される。そして、脈波計測回路にて出力された PPG は、無線通信機により外部の受信機へ送られる。さらに、インターネット回線を通じて遠隔地にあるモニタリング端末へと情報を送ることが可能である。

また、光源には LED (LTSTC170TGKT) を使用し、波長は体動に伴う外乱の影響を受けない 530 nm と設定されている。なお、受光素子はフォトダイオード (TEMD6200FX01)、増幅にはオペアンプ (OPA2330) を用いている。計測回路は 11 mm×11 mm の基板に実装し、紫外線硬化樹脂によって包埋することで防水対策も施されている。これにより、水仕事及び入浴等の日常的な生活環境下での水没の危険性も回避している。

2.4 測定波形

文献 [1] で実施された 24 時間の測定実験で、付け爪型センサから得られた PPG 及び心臓付近に取り付けた電極から得られた ECG を Fig. 2.1 に示す。まず、Fig. 2.1 (a) で示す 24 時間全体の波形を見ると、大部分において波形がクリッピングしていることが分かる。これは、日常的な活動の動作時に生じる（心拍以外の）信号（本論文ではこれを外乱信号と考える）が加算されることによって、センサの測定可能範囲を超えてしまうことが原因である。しかしながら、Fig. 2.1 (b) 左列に示す睡眠時の拡大波形では、PPG 及び ECG のいずれの波形も心拍に由来すると思われる波形（本論文における心拍信号）が測定されている。ただし、ECG はスパイク状の目的信号が現れるのに対し、付け爪型センサで得られる PPG の目的信号は高周波成分をあまり含まない正弦波に近い波形となっている。Fig. 2.1 (b) 中央列の通常活動下や、Fig. 2.1 (c) 右列の激しい運動中では、運動の激しさに応じて外乱信号が増し、心拍に由来する信号の信号対雑音（signal-to-noise: SN）比が極端に低下する。

Table 2.1. Activity scenes in each time period (Data 1)

Japanese standard time	Scene
18:00 – 00:53	Moderate activity
00:53 – 07:00	Sleep
07:00 – 09:47	Moderate activity
09:47 – 11:07	Hard exercise
11:07 – 12:30	Moderate activity
12:30 – 13:56	Sleep
13:56 – 15:09	Hard exercise
15:09 – 18:00	Moderate activity

本論文では、文献 [1] で計測された Fig. 2.1 の信号だけでなく、新たに計測した 2 日分の波形についても心拍推定の調査・解析を行う。これらの計測結果にそれぞれ「データ 1」、「データ 2」、及び「データ 3」とラベルを付け、以後はラベル名を用いて計測波形を指し示す。各データの計測実験に関する詳細は以下の通りである。

データ 1 は 2018 年 5 月 2 日に測定した 24 時間のデータである。各信号のサンプリング周波数は 1 kHz である。このデータは、参考値として胸に取り付けた電極から得られた ECG の信号と、付け爪型センサから得られた PPG の信号の波形が、同じ A/D 変換器に入力されておらず、時間的な同期が取れていない。計測開始のタイミングは ECG と PPG で極力同時としたが、胸部の電極に接続される計測機器と付け爪型センサに接続される計測機器の間でサンプリング周波数はわずかにずれている可能性があり、24 時間という長時間の計測においては、波形後半の時間のずれが顕著となっている可能性がある。また、PPG においては、ダイナミックレンジを超えてクリッピングすることへの対策がされていない測定データである。データ 1 における各時刻での行動は Table 2.1 の通りである。

データ 2 は 2021 年 5 月 3 日に測定した 24 時間のデータである。各信号のサンプリング周波数は 512 Hz である。このデータは、参考値として胸に取り付けた電極から得られた ECG と付け爪型センサから得られた PPG が同じ A/D 変換器に入力されており、時間的な同期が取れている。PPG においては、ダイナミックレンジを超えてクリッピングすることへの対策として、受光センサに遮光テープを取り付けることにより、光の量の調整が施されている測定データである。データ 2 における各時刻での行動は Table 2.2 の通りである。Fig. 2.1 と同様に、データ 2 の波形を Fig. 2.2 に示す。

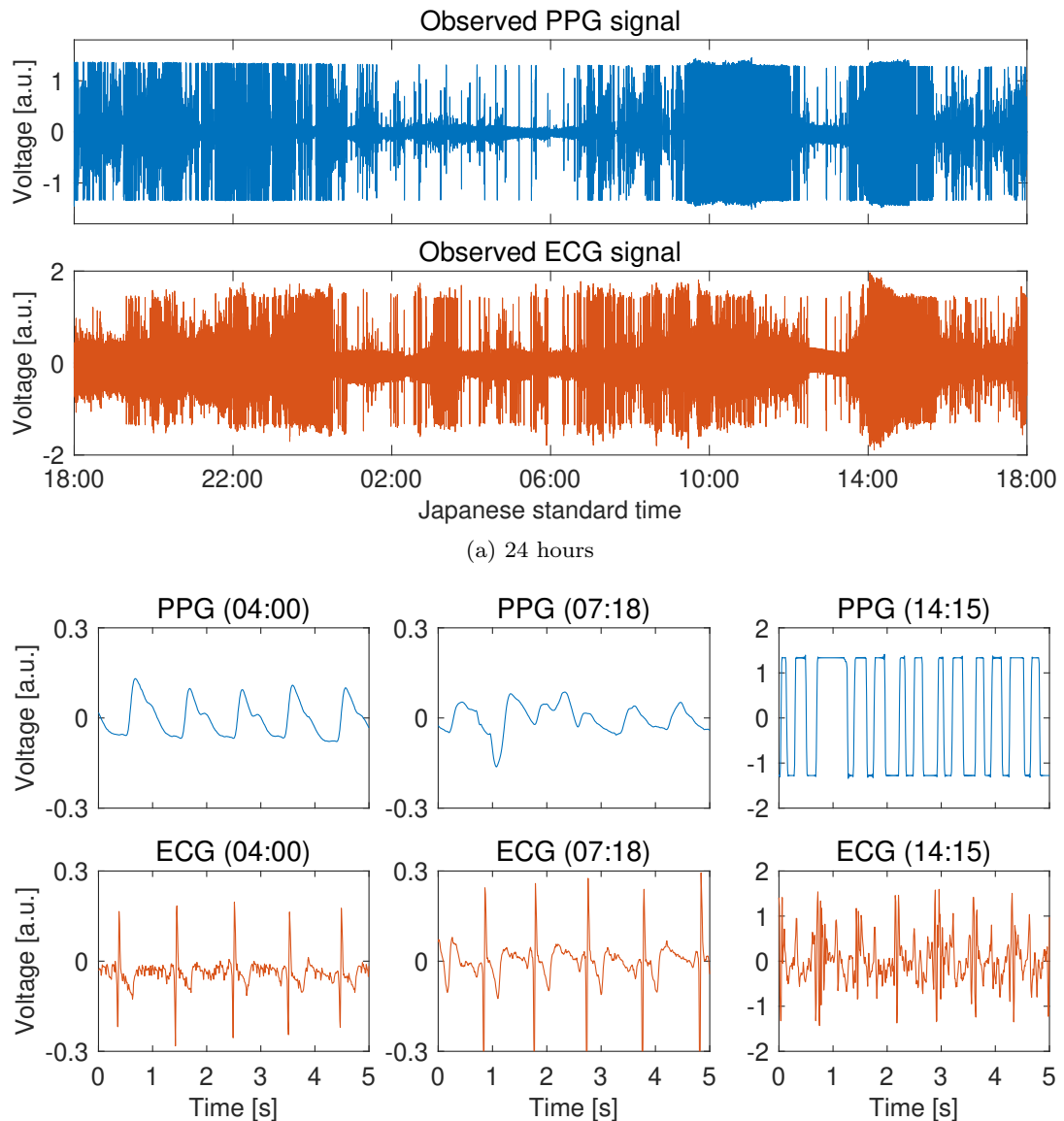
データ 3 は 2021 年 5 月 8 日に測定した 24 時間のデータである。各信号のサンプリング周波数は 1024 Hz である。ECG と PPG の時間同期についてはデータ 2 と同様であるが、データ 3 ではさらに LED にも遮光テープを貼り付け、より減光する処理が施されている。データ 3 における各時刻での行動は Table 2.3 の通りである。Fig. 2.1 と同様に、データ 3 の波形を Fig. 2.3 に示す。

Table 2.2. Activity scenes in each time period (Data 2)

Japanese standard time	Scene
12:00 – 12:50	Moderate activity
12:50 – 13:09	Hard exercise
13:09 – 18:08	Moderate activity
18:08 – 18:30	Hard exercise
18:30 – 21:10	Moderate activity
21:10 – 21:30	Hard exercise
21:30 – 23:03	Moderate activity
23:03 – 07:30	Sleep
07:30 – 10:15	Moderate activity
10:15 – 10:41	Hard exercise
10:41 – 11:20	Moderate activity
11:20 – 11:35	Hard exercise
11:35 – 11:40	Moderate activity
11:40 – 12:09	Hard exercise

Table 2.3. Activity scenes in each time period (Data 3)

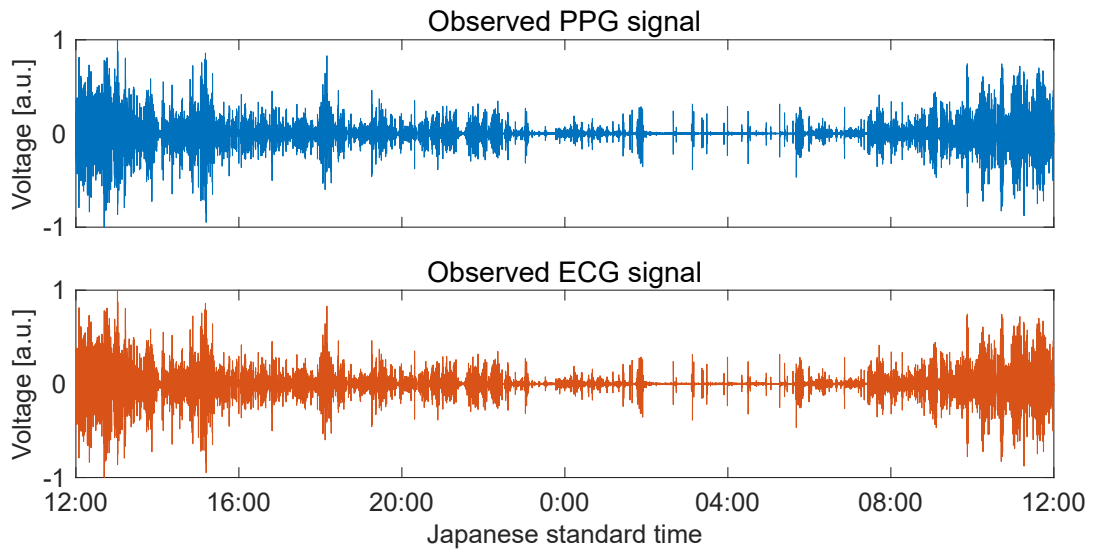
Japanese standard time	Scene
13:00 – 13:11	Moderate activity
13:11 – 13:34	Hard exercise
13:34 – 16:03	Moderate activity
16:03 – 16:17	Hard exercise
16:17 – 17:45	Moderate activity
17:45 – 18:05	Hard exercise
18:05 – 21:46	Moderate activity
21:46 – 22:05	Hard exercise
22:05 – 23:50	Moderate activity
23:50 – 06:10	Sleep
06:10 – 09:08	Moderate activity
09:08 – 09:53	Hard exercise
09:53 – 11:34	Moderate activity
11:34 – 11:48	Hard exercise
11:48 – 12:15	Moderate activity
12:15 – 12:32	Hard exercise
12:32 – 13:05	Moderate activity



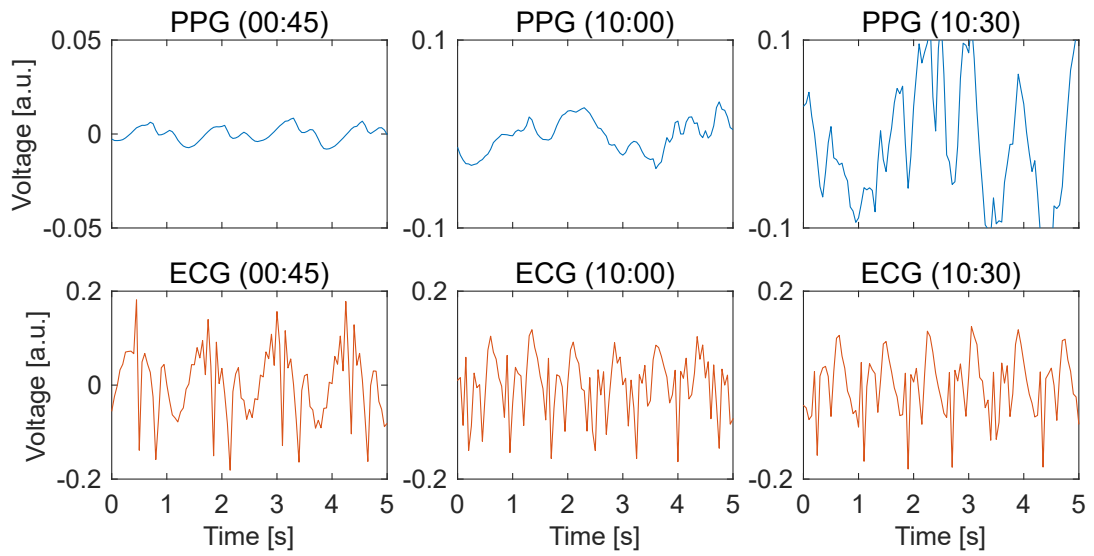
(b) Five seconds in sleep (left), moderate activity (center), and hard exercise (right) periods
 Fig. 2.1. PPG (top) and ECG (bottom) signals in Data 1 for (a) 24 hours and (b) five seconds.

2.5 本章のまとめ

本章では、付け爪型センサの概要について詳しく解説し、従来の胸部接触型の電極センサに対する利点を述べた。さらに、従来の胸部接触型の電極センサから得られる ECG の波形と比較したとき、付け爪型センサから得られる PPG では、外乱によって波形が大きく乱れるという問題点を述べた。次章では、ノイズの多く含まれる付け爪型センサの PPG から高精度に心拍を推定するため、音楽信号解析の分野における基本周波数推定法の適用を検討する。そし



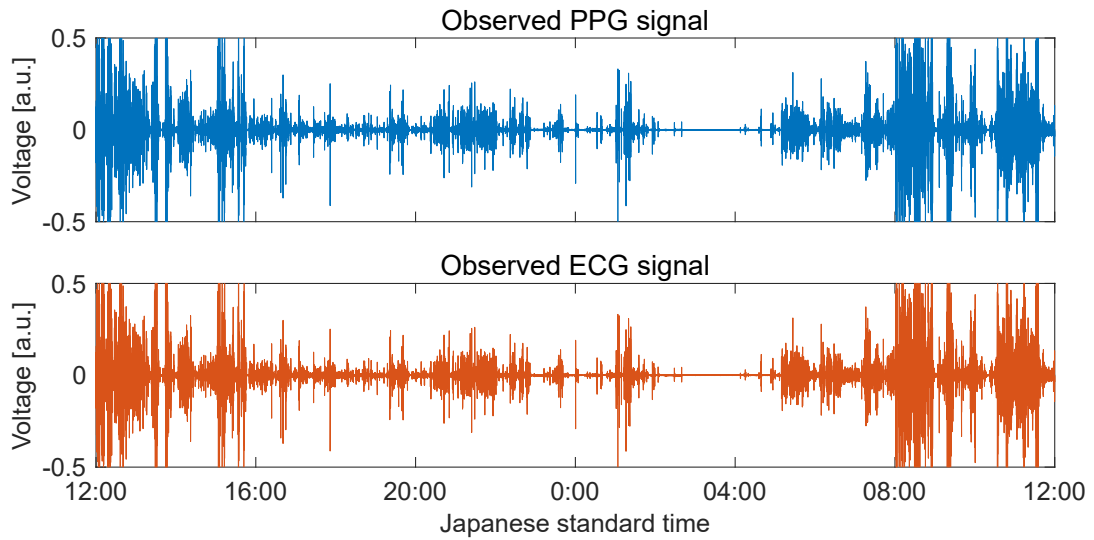
(a) 24 hours



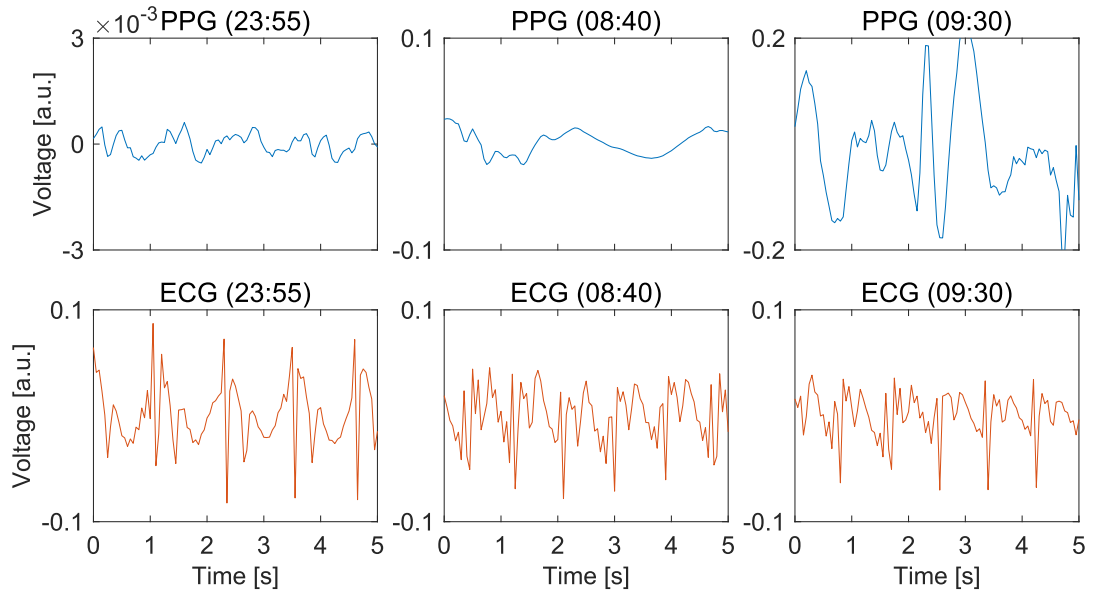
(b) Five seconds in sleep (left), moderate activity (center), and hard exercise (right) periods

Fig. 2.2. PPG (top) and ECG (bottom) signals in Data 2 for (a) 24 hours and (b) five seconds.

て、代表的な基本周波数推定法の理論について説明する。



(a) 24 hours



(b) Five seconds in sleep (left), moderate activity (center), and hard exercise (right) periods

Fig. 2.3. PPG (top) and ECG (bottom) signals in Data 3 for (a) 24 hours and (b) five seconds.

第 3 章

音楽信号解析における代表的な基本周波数推定法

3.1 まえがき

本章では、本研究で用いる音楽信号解析分野における代表的な 4 手法について、それぞれ原理を説明する。まず 3.2 節では、付け爪型センサから得られる PPG に対して、音楽信号解析分野の基本周波数推定法を適用する動機について述べる。さらに、各基本周波数推定法の概要をそれぞれ述べる。3.3 節では、以降の節で解説する基本周波数推定法を用いた信号の解析に必要な前処理について説明する。3.4 節では、正規化自己相関関数法に基づく心拍推定について説明する。3.5 節では、ケプストラム法に基づく心拍推定について説明する。3.6 節では、YIN 法に基づく心拍推定について説明する。3.7 節では、MUSIC 法に基づく心拍推定について説明する。3.8 節では、本章のまとめを述べる。

3.2 本研究の動機

生体信号処理の分野において対象となる信号は、安静な（ノイズの入りづらい）状態での測定が基本とされており、Fig. ?? (b) のようなスパイクがはっきりと確認できるような波形を扱うことが多い。従って、生体信号処理の分野では、測定で得られた信号におけるノイズは問題となりづらく、RRI に基づく基本周期推定によって心拍が高精度に推定できる場合がほとんどである。しかし、本研究で使用する付け爪型センサ [1] から得られる PPG は、体動等に起因する大きなノイズが問題となっており、ゼロクロス法等の原始的な心拍推定法では推定が困難である。

そこで本研究では、音楽信号解析分野の基本周波数推定法を用いることで、PP に対する高精度な心拍推定を目的とする。特に、正規化自己相関関数法、ケプストラム法、YIN 法、及び MUSIC 法の 4 手法を音楽信号解析における代表的な基本周波数推定法として取り上げ、付け爪型センサで得られる PPG に適用することを試みる。正規化自己相関関数法は、信号処理に

における最も基本的な基本周波数推定法であり、Wiener-Khinchin の定理より、信号のパワースペクトルから基本周波数を推定していることと等価である。ケプストラム法は、ボーカルや音声信号の基本周波数を推定するために提案された手法であり、信号の対数振幅スペクトルの周波数成分を解析して基本周波数を得る。YIN 法は、正規化自己相関関数で用いられる自己相関関数を二乗誤差に置き換えた特徴量を用いており、よりノイズに頑健な基本周波数推定法として提案された技術である。最後に、MUSIC 法は、信号と無相関なノイズが混入した観測を仮定し、部分空間法により信号空間と雑音空間を分離することで信号の基本周波数を推定する手法である。

3.3 信号の前処理

観測した PPG を $(x_{\text{raw}}[t])_{t=1}^T$ とする。ここで、 $t = 1, 2, \dots, T$ は離散時間インデックスであり、 T は観測時間長である。PPG として測定される電圧波形には直流成分が含まれている。基本周波数を解析する上では、この直流成分を事前に除去しておくことが望ましいため、前処理としてハイパスフィルタを適用する。本研究では、Butterworth IIR デジタルフィルタを用いて、信号 $(x[t])_{t=1}^T$ に変換する。この時、カットオフ周波数は 0.3 Hz 及びフィルタ次数は 30 次を設定している。このフィルタの振幅応答と位相応答を Fig. 3.1 に示す。

次に、 $(x[t])_{t=1}^T$ を短時間区間信号に分割する。短時間区間の信号長を L 、短時間区間のシフト長を S とすると、 $(x[t])_{t=1}^T$ の短時間区間信号への分割は次式で表される。

$$x_s[i, l] = x[l + (i - 1)S] \quad (3.1)$$

ここで、 $i = 1, 2, \dots, I$ は短時間区間インデックス、 $l = 1, 2, \dots, L$ は短時間区間内の離散時間インデックスをそれぞれ表す。この操作を Fig. 3.2 に示す。

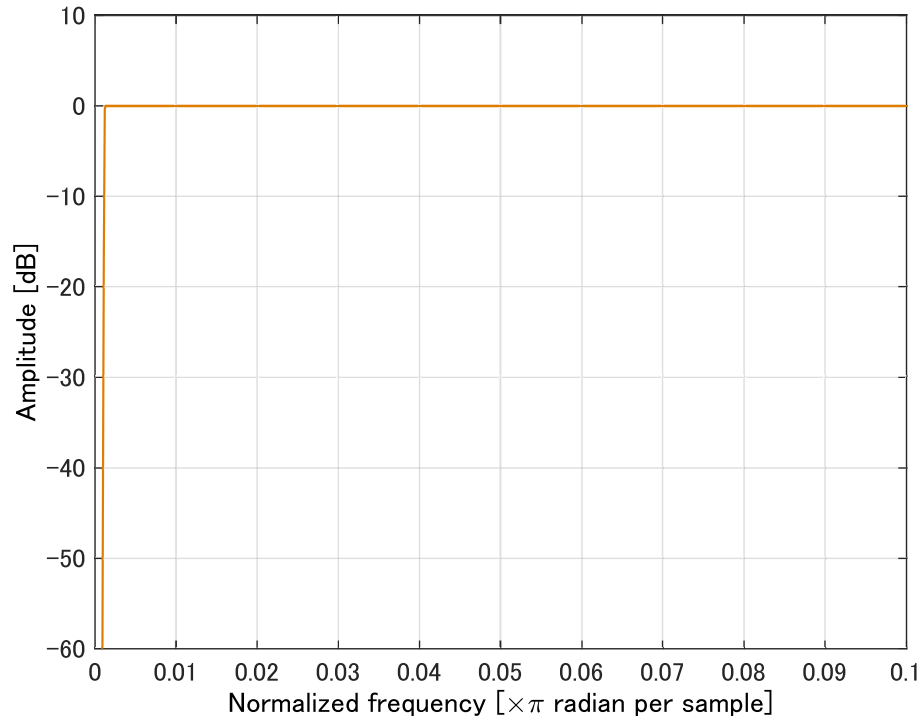
3.4 正規化自己相関関数法に基づく心拍推定

正規化自己相関関数法 [2] は、正規化された自己相関関数を用いて信号の基本周期を求める手法である。自己相関関数とは、波形を時間シフトさせる前後で相関係数を算出したものであり、時間シフト量 τ に対する関数となる。本節では、正規化自己相関関数法の推定手法について示す。

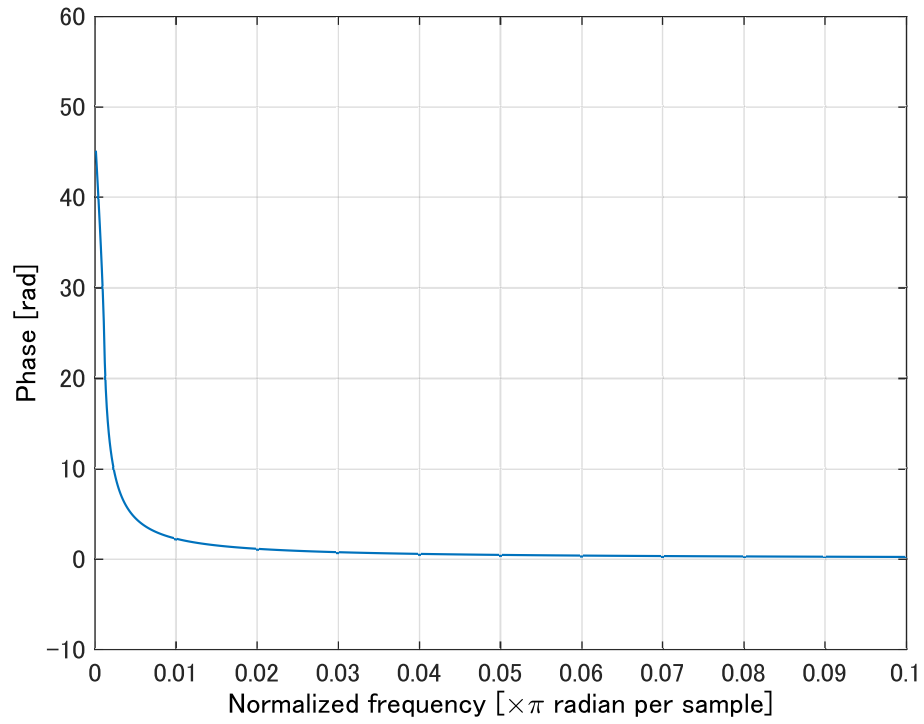
まず、正規化された自己相関関数 $\beta[i, \tau]$ を定義する。これは、短時間区間の時間波形 $x_s[i, l]$ に対して、時間方向に τ だけずらした信号 $x_s[i, l + \tau]$ を掛け合わせ、時間方向に平均することにより求まる。ただし、 $\tau = 0$ の時、1 となるように正規化されている。したがって、正規化された自己相関関数 $\beta[i, \tau]$ は次式となる。

$$\beta[i, \tau] = \frac{1}{\sum_{l=1}^L x_s[i, l]^2} \sum_{l=1}^{L-\tau} x_s[i, l] x_s[i, l + \tau] \quad (3.2)$$

正弦波のように完全な周期では、 τ がちょうど波形の 1 周期分（又は 1 周期の整数倍分）と



(a) Amplitude



(b) Phase

Fig. 3.1. Characteristic of high-pass filter used in PPG and ECG preprocessing: (a) amplitude and (b) phase responses.

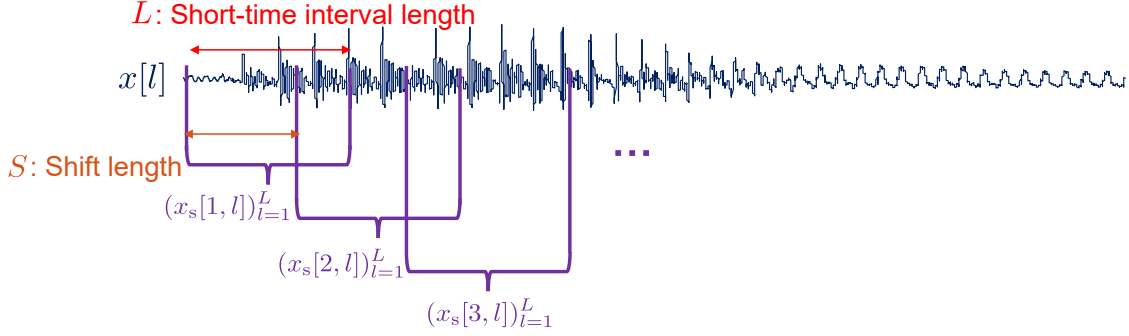


Fig. 3.2. Signal splitting into short-time signals.

なるように時間シフトすると、正規化された自己相関関数は1となる。短時間区間に分割された心拍由来のPPGは、ノイズがなければほとんど周期波形とみなせるため、正規化された自己相関関数 $\beta[i, \tau]$ にはシフト量 τ に応じて複数個のピークが現れる。従って、正規化された自己相関関数の最初のピーク点のシフト量が短時間区間毎の基本周期の推定値 $(\hat{\tau}_o[i])_{i=1}^I$ となり、その逆数が推定基本周波数（推定心拍） $(\hat{f}_o[i])_{i=1}^I$ となる。この処理は次のように表される。

$$\hat{\tau}_o[i] = \arg \max_{\tau \in [\tau_{\min}, \tau_{\max}]} \beta[i, \tau] \quad (3.3)$$

$$\hat{f}_o[i] = \frac{f_s}{\tau_o[i]} \quad (3.4)$$

ここで、 f_s は信号のサンプリング周波数である。式 (3.3) の最大値の探索は、基本周波数の推定範囲の下限 $f_{\min}$ と上限 $f_{\max}$ を事前に定めたうえで、それらに対応するシフト量の下限 $\tau_{\min} = f_s / f_{\max}$ 及び上限 $\tau_{\max} = f_s / f_{\min}$ の範囲内で行う。すなわち、 $\tau_{\min}$ 及び $\tau_{\max}$ は、自己相関関数の最初のピーク点があると思われる範囲を表すための（シフト量の）始点と終点である。この一連の処理を Fig. 3.3 に示す。なお、心拍は一般的に1分間当たりの拍数である beats per minute (bpm) で表現されるため、推定心拍は $\hat{h}_o[i] = 60\hat{f}_o[i]$ となる。

3.5 ケプストラム法に基づく心拍推定

ケプストラムとは、信号の対数振幅スペクトルの周波数成分を解析したものであり、音声信号処理の分野では基本周波数の推定に古くから用いられてきた歴史がある。ケプストラム法 [3] では、まず $(x_s[i, l])_{l=1}^L$ に何らかの窓関数 $(w[l])_{l=1}^L$ を乗じて L 点離散 Fourier 変換 (discrete Fourier transform: DFT) を施し、短時間区間の複素スペクトル $(F[i, \omega])_{\omega=1}^{\Omega}$ を得る。

$$F[i, \omega] = \sum_{l=1}^L w[l] x_s[i, l] \exp \left[-j \frac{2\pi(\omega-1)(l-1)}{L} \right] \quad (3.5)$$

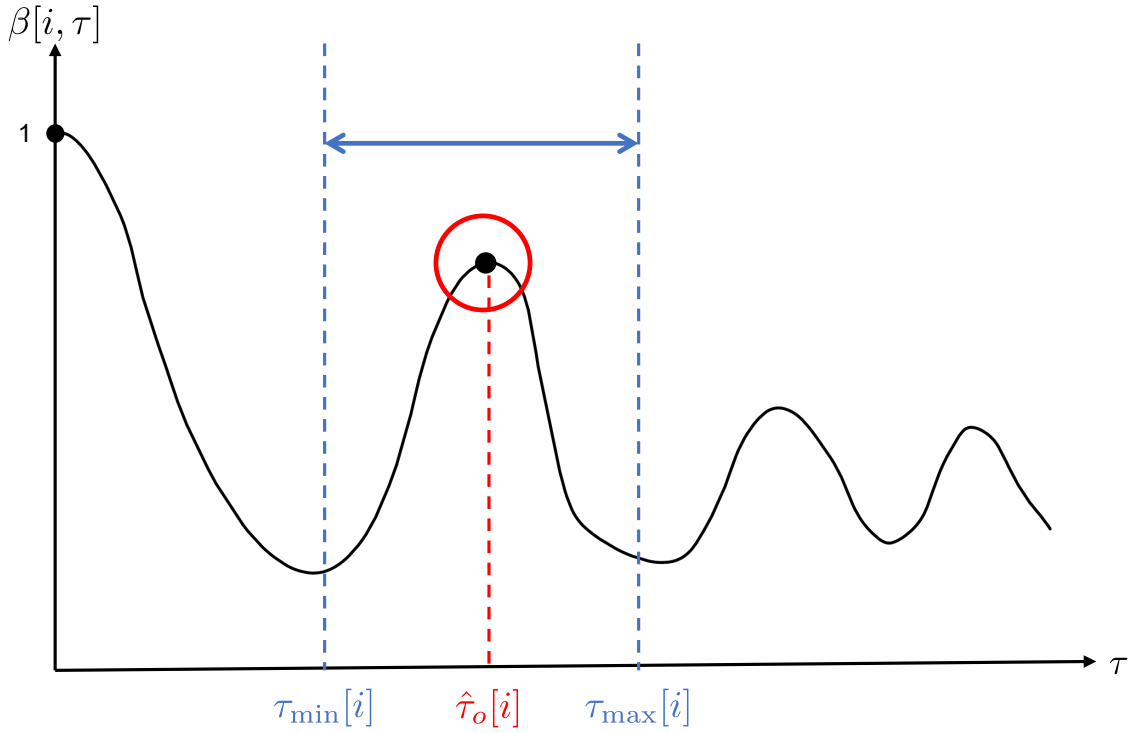


Fig. 3.3. Estimation of fundamental period in short time interval based on normalized autocorrelation function.

ここで, $j = \sqrt{-1}$ であり, $\omega = 1, 2, \dots, \Omega$ は離散周波数インデックスである (但し $\Omega = L$). 実数ケプストラム $(C[i, q])_{q=1}^Q$ は $(F[i, \omega])_{\omega=1}^{\Omega}$ の絶対値の対数を逆 DFT して得られる.

$$C[i, q] = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \log |F[i, \omega]| \exp \left[j \frac{2\pi(\omega-1)(l-1)}{L} \right] \quad (3.6)$$

ここで, $q = 1, 2, \dots, Q$ はケフレンシと呼ばれるケプストラム領域のインデックスである (但し $Q = L$). 即ち, ケプストラム $(C[i, q])_{q=1}^Q$ は対数スペクトルのスペクトル^{*1}であり, 基本周波数 $f_o[i]$ とその倍音 $(nf_o[i])_{n=2}^N$ の周波数をもつ正弦波から成る目的信号のスペクトル解析ができるため, 基本周波数推定に用いられる. なお, ケプストラム $(C[i, q])_{q=1}^Q$ の時間変動はケプストログラムと呼ばれる.

短時間区間毎の基本周波数の推定値 $(\hat{f}_o[i])_{i=1}^I$ は, 次式のようにケプストラム $(C[i, q])_{q=1}^Q$ が最大値を取るケフレンシインデックス $(\hat{q}_o[i])_{i=1}^I$ から求められる.

$$\hat{q}_o[i] = \arg \max_{q \in [q_{\min}, q_{\max}]} C[i, q] \quad (3.7)$$

$$\hat{f}_o[i] = \frac{f_s}{\hat{q}_o[i] - 1} \quad (3.8)$$

^{*1} より厳密には, $\log |F[i, \omega]|$ は ω に関して偶関数であるため, 実数ケプストラム $C[i, q]$ は $\log |F[i, \omega]|$ の cos 展開係数といえる.

式 (3.7) の最大値の探索は、基本周波数の推定範囲の下限 $f_{\min}$ と上限 $f_{\max}$ を事前に定めたいえで、それらに対応するケフレンシインデックスの下限 $q_{\min} = \text{round}(f_s/f_{\max} + 1)$ 及び上限 $q_{\max} = \text{round}(f_s/f_{\min} + 1)$ の範囲内で行う。ここで、 $\text{round}(\cdot)$ は整数への丸め処理を表す。

3.6 YIN 法に基づく心拍推定

YIN 法 [4] は話者やボーカル等、人間の音声の基本周波数を頑健に推定するために提案された手法である。この手法では、波形を時間シフトさせる前後で二乗誤差を算出し、時間シフト量に対する誤差関数を定義する。すなわち、NCF 法において、式 (3.2) で定義される自己相関関数中の内積を二乗誤差に置き換えたものと解釈できる。

信号 $x_s[i, l]$ に対し、シフト量 τ だけずらした信号 $x_s[i, l + \tau]$ との二乗誤差関数は次式のよに定義される。

$$\delta[i, \tau] = \sum_{l=1}^{L/2} (x_s[i, l] - x_s[i, l + \tau])^2 \quad (3.9)$$

基本周期に対応する時間サンプル $\tau_o = f_s/f_o$ において、信号 $x_s[i, l]$ が周期的である場合、式 (3.9) はシフト量が基本周期に一致する ($\tau = \tau_o$) 時に小さい値を示す。

次に、この誤差関数を、次式に示すように累積平均で正規化する。

$$\delta'[i, \tau] = \frac{\delta[i, \tau]}{\frac{1}{\tau} \sum_{\tau'=1}^{\tau} \delta[i, \tau']} \quad (3.10)$$

最後に、 $\delta'[i, \tau]$ に対して、ある閾値 α 以下となるシフト量 τ を求め、その中で最も小さい τ が推定基本周期に対応する時間サンプル $\hat{\tau}_o[i]$ として得られる。すなわち、集合 $\{\tau \in \mathbb{N} \mid \alpha \geq \delta'[i, \tau]\}$ の中で最小の τ が $\hat{\tau}_o[i]$ である。また、推定基本周波数は $\hat{f}_o[i] = f_s/\hat{\tau}_o[i]$ として得られる。この一連の処理を Fig. 3.4 に示す。なお、閾値 α は（音声信号処理においては）通常 $\alpha = 0.10 \sim 0.15$ に設定される。

3.7 MUSIC 法に基づく心拍推定

MUSIC 法 [5] は部分空間法に基づく信号の分類手法であり、センサアレイ信号処理で信号の到来方向推定によく用いられる。一方で、時間信号に MUSIC 法を適用し基本周波数を推定する手法 [8, 9] も提案されており、信号とは無相関なノイズの混入に対して頑健という特徴がある。

まず、短時間区間信号 $(x_s[i, l])_{l=1}^L$ 中の長さ K （但し $K < L$ ）の部分信号を切り出して次式

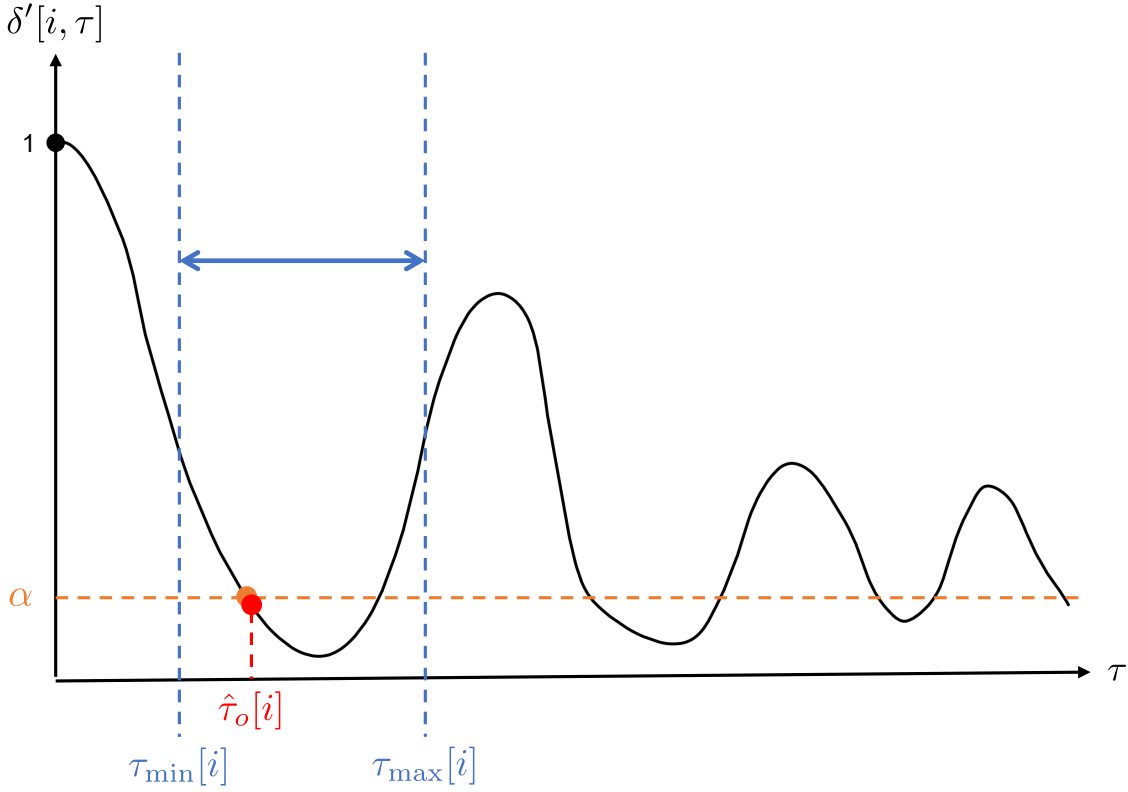


Fig. 3.4. Detection of time lag that corresponds to fundamental period in short time interval based on YIN method.

のようにベクトルで定義する.

$$\mathbf{x}_{i,m} = \begin{bmatrix} x_s[i, 1 + (m-1)P] \\ \vdots \\ x_s[i, k + (m-1)P] \\ \vdots \\ x_s[i, K + (m-1)P] \end{bmatrix} \in \mathbb{R}^K \quad (3.11)$$

ここで, $m = 1, 2, \dots, M$ は $x_s[i, l]$ 中の部分信号インデックス, $k = 1, 2, \dots, K$ は部分信号内の離散時間インデックス, P は部分信号を切り出す際のシフト長をそれぞれ表す. このベクトルより, i 番目の短時間区間の信号相関行列 $\mathbf{R}_i$ が次式のように得られる.

$$\begin{aligned} \mathbf{R}_i &= \mathbb{E}[\mathbf{x}_{i,m} \mathbf{x}_{i,m}^T] \\ &\approx \frac{1}{M} \sum_{m=1}^M \mathbf{x}_{i,m} \mathbf{x}_{i,m}^T \in \mathbb{R}^{K \times K} \end{aligned} \quad (3.12)$$

ここで, $\cdot^T$ は転置, $\mathbb{E}[\cdot]$ を表す. この操作を Fig. 3.5 に示す. MUSIC 法では, 信号相関行列 $\mathbf{R}_i$ の部分空間から目的信号成分とノイズ成分を分離し, MUSIC スペクトルと呼ばれる特徴

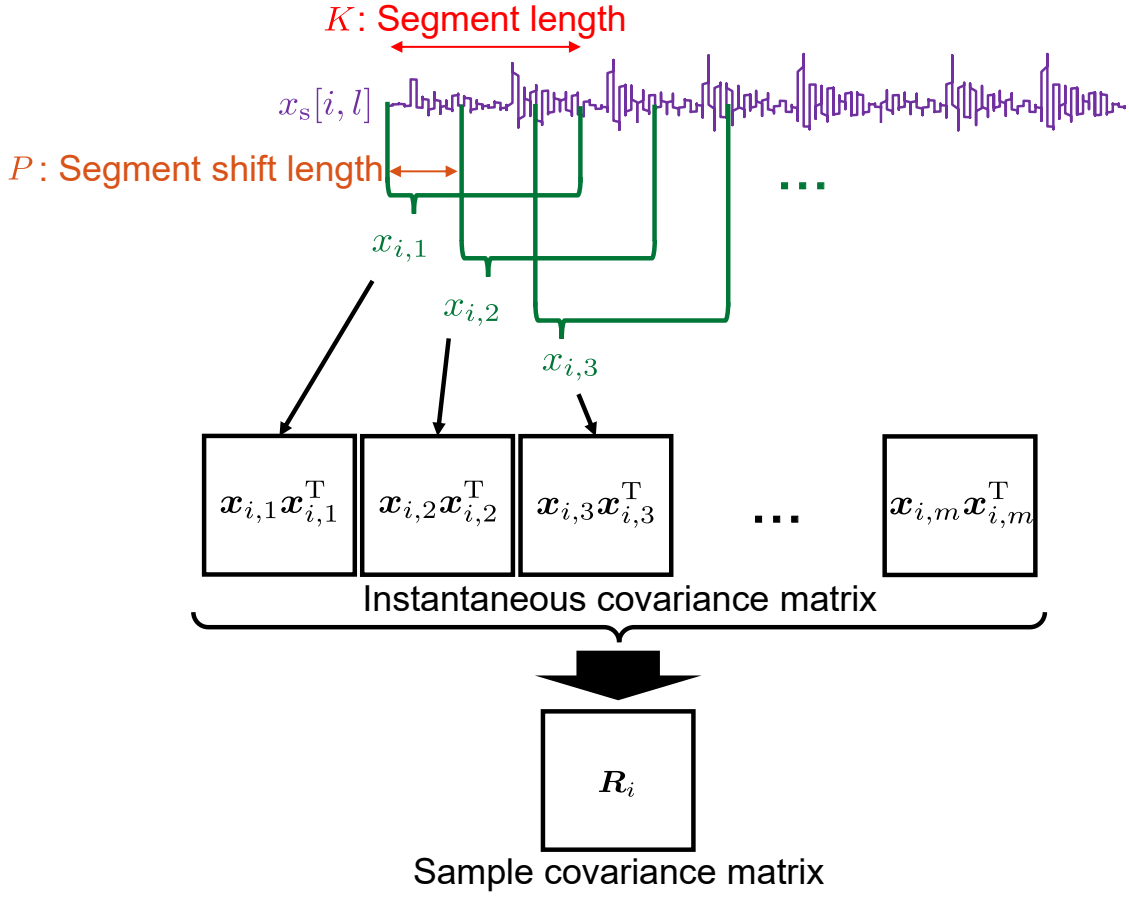


Fig. 3.5. Heart rate estimation algorithm in MUSIC method.

量を算出する。信号相関行列 R_i を固有値分解すると、次式となる。

$$R_i = V_i \Lambda_i V_i^{-1} \quad (3.13)$$

$$\Lambda_i = \begin{bmatrix} \lambda_{i,1} & 0 & \cdots & 0 \\ 0 & \lambda_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{i,K} \end{bmatrix} \quad (3.14)$$

$$V_i = [v_{i,1} \quad v_{i,2} \quad \cdots \quad v_{i,K}] \quad (3.15)$$

ここで、 $\lambda_{i,k}$ は R_i の固有値であり、 $\lambda_{i,1} \geq \lambda_{i,2} \geq \cdots \geq \lambda_{i,K}$ のように降順に並んでいるものとする。また、 $v_{i,k}$ は $\lambda_{i,k}$ に対応する固有ベクトルである。いま、短時間区間信号 $(x_s[i, l])_{l=1}^L$ が次式のように表されると仮定する。

$$x_s[i, l] = s[i, l] + \rho[i, l] \quad (3.16)$$

$$s[i, l] = \sum_{n=1}^N A_n[i] \sin(2\pi n f_o[i] l + \theta_n[i]) \quad (3.17)$$

ここで、 $A_n[i]$ 及び $\theta_n[i]$ は n 次基本成分の振幅及び位相差をそれぞれ表し、 $(\rho[i, l])_{l=1}^L$ は

$(s[i, l])_{l=1}^L$ と無相関なノイズ信号である。即ち、観測信号は基本周波数 $f_o[i]$ とその倍音 $(nf_o[i])_{n=2}^N$ の周波数をもつ正弦波から成る目的信号 $(s[i, l])_{l=1}^L$ とノイズ信号 $(\rho[i, l])_{l=1}^L$ の混合と仮定する。もし各信号のパワー比が $\sum_{l=1}^L |s[i, l]|^2 \gg \sum_{l=1}^L |\rho[i, l]|^2$ であれば、式 (3.13) で得られる固有値のうち、上位 $2N$ 番目までの固有値 $(\lambda_{i,n})_{n=1}^{2N}$ に対応する固有ベクトル $(\mathbf{v}_{i,n})_{n=1}^{2N}$ が目的信号 $(s[i, l])_{l=1}^L$ の部分空間を張り、残りの固有値 $(\lambda_{i,n})_{n=2N+1}^K$ に対応する固有ベクトル $(\mathbf{v}_{i,n})_{n=2N+1}^K$ がノイズ信号 $(\rho[i, l])_{l=1}^L$ の部分空間を構成する*2。従って、式 (3.13) は次式のように分解できる。

$$\mathbf{R}_i = \mathbf{U}_i^{(s)} \mathbf{D}_i^{(s)} \mathbf{U}_i^{(s)-1} + \mathbf{U}_i^{(\rho)} \mathbf{D}_i^{(\rho)} \mathbf{U}_i^{(\rho)-1} \quad (3.18)$$

$$\mathbf{D}_i^{(s)} = \begin{bmatrix} \lambda_{i,1} & 0 & \cdots & 0 \\ 0 & \lambda_{i,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{i,2N} \end{bmatrix} \quad (3.19)$$

$$\mathbf{U}_i^{(s)} = [\mathbf{u}_{i,1} \quad \mathbf{u}_{i,2} \quad \cdots \quad \mathbf{u}_{i,2N}] \quad (3.20)$$

$$\mathbf{D}_i^{(\rho)} = \begin{bmatrix} \lambda_{i,2N+1} & 0 & \cdots & 0 \\ 0 & \lambda_{i,2N+2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{i,K} \end{bmatrix} \quad (3.21)$$

$$\mathbf{U}_i^{(\rho)} = [\mathbf{u}_{i,2N+1} \quad \mathbf{u}_{i,2N+2} \quad \cdots \quad \mathbf{u}_{i,K}] \quad (3.22)$$

理想的には、式 (3.18) の第一項及び第二項がそれぞれ目的信号の信号相関行列及びノイズ信号の信号相関行列に対応する。

MUSIC スペクトルは、ノイズ信号の固有ベクトル $(\mathbf{u}_{i,k})_{k=2N+1}^K$ ($\mathbf{U}_i^{(\rho)}$ の列ベクトル) を用いて次式で計算される。

$$G[i, \gamma] = \frac{1}{\sum_{k=2N+1}^K |\mathbf{u}_{i,k}^T \mathbf{b}[\gamma]|^2} \quad (3.23)$$

$$\mathbf{b}[\gamma] = \begin{bmatrix} 1 \\ e^{-j \frac{2\pi(\gamma-1)}{K}} \\ e^{-j \frac{4\pi(\gamma-1)}{K}} \\ \vdots \\ e^{-j \frac{2\pi(K-1)(\gamma-1)}{K}} \end{bmatrix} \in \mathbb{C}^K \quad (3.24)$$

ここで、 $\gamma = 1, 2, \dots, \Gamma$ は離散周波数インデックスである (但し $\Gamma = K$)。 $\mathbf{b}[\gamma]$ は DFT 基底であるため、 $\mathbf{u}_{i,k}^T \mathbf{b}[\gamma]$ はノイズ信号の固有ベクトル $\mathbf{u}_{i,k}$ の DFT そのものである。式 (3.23) はノイズ信号の固有ベクトルの DFT が分母にあることから、MUSIC スペクトル $(G[i, \gamma])_{\gamma=1}^\Gamma$ は、目的信号 $(s[i, l])_{l=1}^L$ の部分空間に対応する周波数 (式 (3.17) における正弦波周波数 $(nf_o)_{n=1}^N$) で顕著なピークを持つ。従って、ノイズ信号に対して頑健に基本周波数の推定が可能となる。

*2 目的信号に対応する固有値の数が正弦波の数 N の 2 倍となるのは、式 (3.17) において実数正弦波の重ね合わせを仮定しているためである。複素正弦波の重ね合わせであれば、対応する固有値の数と正弦波個数は一致する [9]。

短時間区間毎の基本周波数の推定値 $(\hat{f}_o[i])_{i=1}^I$ は、次式のように MUSIC スペクトル $(G[i, \gamma])_{\gamma=1}^{\Gamma}$ が最大値を取る周波数インデクス $(\hat{\gamma}_o[i])_{i=1}^I$ から求められる.

$$\hat{\gamma}_o[i] = \arg \max_{\gamma \in [\gamma_{\min}, \gamma_{\max}]} G[i, \gamma] \quad (3.25)$$

$$\hat{f}_o[i] = \frac{f_s}{K} (\hat{\gamma}_o[i] - 1) \quad (3.26)$$

NCF 法やケプストラム法の場合と同様に、実際は推定範囲の下限 $f_{\min}$ と上限 $f_{\max}$ に対応する離散周波数インデクスの下限 $\gamma_{\min} = \text{round}(K f_{\min} / f_s + 1)$ 及び上限 $\gamma_{\max} = \text{round}(K f_{\max} / f_s + 1)$ の範囲内で式 (3.25) の最大値探索を行う.

3.8 本章のまとめ

本章では、生体信号処理分野において、ノイズに対して頑健な心拍推定手法の検討があまりなされていないことについて触れ、音楽信号解析分野の周波数推定法の適用可能性について述べた. さらに、本論文で実際に適用する正規化自己相関関数法、ケプストラム法、YIN 法、及び MUSIC 法の 4 手法についてそれぞれ数式を用いて詳しく説明した. 次章では、本章で解説した 4 手法を実際に ECG 及び PPG に適用し、推定精度の比較を行う.

第 4 章

実験

4.1 まえがき

本章では、前章で解説した 4 手法を実際の PPG に適用し、基本周波数推定の精度を比較する。4.2 節では、精度比較のために用いる客観的評価指標についてそれぞれ解説する。4.3 節では、データ 1 に各手法を適用した際の結果について述べる。但しデータ 1 は、2.4 節で述べた通り、心臓付近に貼り付けた接触型心拍センサによる ECG（心拍の参考値）と、付け爪型センサによる PPG の時間同期がとれていないため、データ 1 に対する実験結果は参考に留める点に注意する。さらに、4.4 節及び 4.5 節にて、データ 2 及びデータ 3 に各手法を適用した際の結果について述べる。これらのデータは、ECG 及び PPG の時間同期が取れているほか、異なる度合いで PPG のクリッピング対策が施されている。4.6 節では、本章のまとめを述べる。本章では以後、正規化自己相関関数法を NCF、ケプストラム法を CEP、YIN 法を YIN、及び MUSIC 法を MUS という略語で示す。

4.2 評価指標

本実験では、3 章に示す 4 手法を PPG に適用して得られる推定心拍 $\hat{h}_o[i]$ の精度について実験的に調査する。このとき、心拍の参考値（正解値）は、同時に計測された ECG に対して、医療分野で一般的である RRI 推定を適用し算出される心拍値とし、これを $h_o[i]$ と記述する。また、推定精度の評価指標には、平均二乗誤差（mean squared error: MSE）、平方平均二乗誤差（root mean squared error: RMS）、グロスピッチ誤差（gross pitch error: GPE）[10, 11]、及びファインピッチ誤差（fine pitch error: FPE）[10, 11] の 4 つを用いた。それぞれの評価指標の定義は次の通りである。

- MSE

MSE は、正解値と推定値の二乗誤差を計算し、その平均を評価値として与える評価指

標であり，次式で表される．

$$\text{MSE} = \frac{1}{I} \sum_{i=1}^I \left(h_o[i] - \hat{h}_o[i] \right)^2 \quad (4.1)$$

- RMS

RMS は，正解値と推定値の二乗誤差を検出し，その平方平均を評価値として与える評価指標である．即ち，MSE の平方根であり，次式で表される．

$$\text{RMS} = \sqrt{\frac{1}{I} \sum_{i=1}^I \left(h_o[i] - \hat{h}_o[i] \right)^2} \quad (4.2)$$

- GPE

GPE は， $\hat{h}_o[i]$ と $h_o[i]$ の絶対誤差が $p \times h_o[i]$ を上回る短時間区間の割合を算出したものである．すなわち，集合 $\mathbb{I} = \{i \in \mathbb{N} \mid p \times h_o[i] < |\hat{h}_o[i] - h_o[i]|\}$ を用いて，次式で表される．

$$\text{GPE} = \frac{1}{I} |\mathbb{I}| \quad [\%] \quad (4.3)$$

例えば，GPE が 0 % のとき，すべての短時間区間において，誤差率を p % と定義すると， $\hat{h}_o[i]$ と $h_o[i]$ の絶対誤差が $p \times h_o[i]$ 以下であり，誤差率 p % 以内で推定できていることを示す．したがって，GPE は大まかな推定精度を表す評価指標である．

- FPE

FPE は， $\hat{h}_o[i]$ と $h_o[i]$ の絶対誤差が $p \times h_o[i]$ を下回る短時間区間（GPE において正解とみなす短時間区間）のみを用いて，絶対誤差の標準偏差を算出したものである．すなわち，次式で表される．

$$\text{FPE} = \sqrt{\frac{1}{|\bar{\mathbb{I}}|} \sum_{i \in \bar{\mathbb{I}}} \left\{ (\hat{h}_o[i] - h_o[i])^2 - \left[\frac{1}{|\bar{\mathbb{I}}|} \sum_{i \in \bar{\mathbb{I}}} (\hat{h}_o[i] - h_o[i]) \right]^2 \right\}} \quad (4.4)$$

ここで， $\bar{\mathbb{I}}$ は集合 $\mathbb{I}$ の補集合（すなわち， $\bar{\mathbb{I}} = \{i \in \mathbb{N} \mid p \times h_o[i] \geq |\hat{h}_o[i] - h_o[i]|\}$ ）であり，GPE において正解とみなす短時間区間の集合）である FPE は絶対誤差の標準偏差であるため，0 に近いほど高精度な推定ができていることを示す．したがって，FPE は大まかに推定できている短時間区間における細部の推定精度を表す評価指標である．

4.3 データ 1 に対する実験結果

4.3.1 実験条件

心拍推定においては，PPG のサンプリング周波数 f_s を測定時の 1 kHz から 20 Hz にダウンサンプリングして解析した．本論文で用いたダウンサンプリングは全て下記の手順となっており，これはディジタル信号処理における一般的なダウンサンプリング法である．

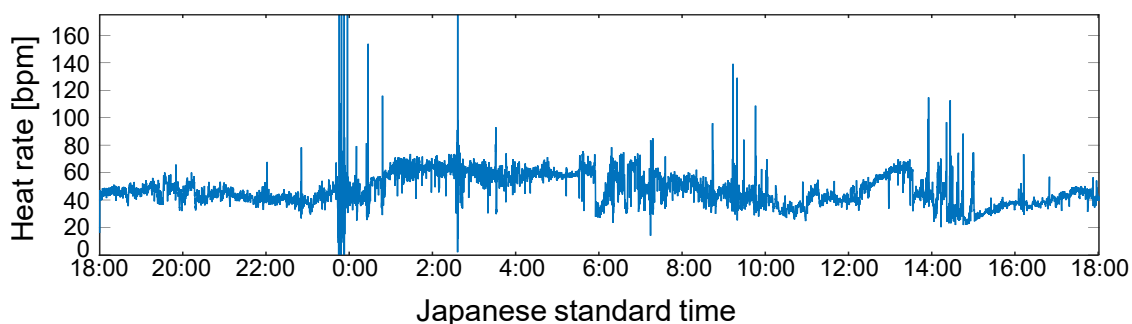


Fig. 4.1. Calculated heart rate obtained from ECG signal (Data 1) resampled at 78.125 mHz and used as reference.

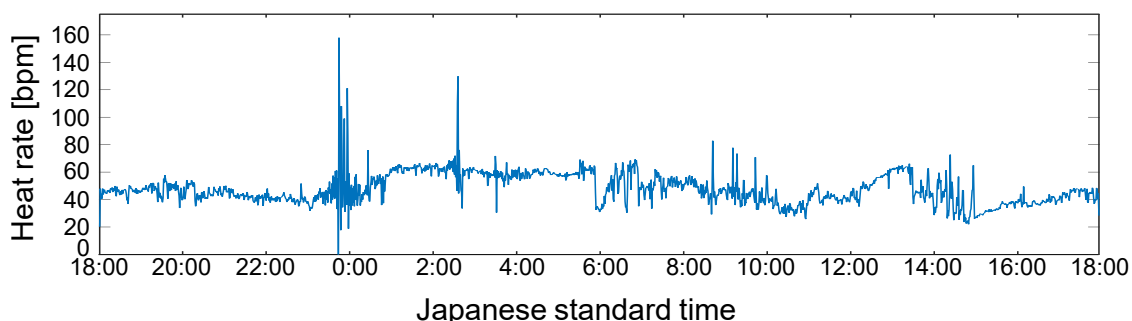


Fig. 4.2. Calculated heart rate obtained from ECG signal (Data 1) resampled at 19.53125 mHz and used as reference.

1. エイリアシングによる歪みを避けるため、ダウンサンプリング前の信号に対して FIR ローパスフィルタ（フィルタ次数は 1000 点、カットオフ周波数はダウンサンプリング後のナイキスト周波数）を適用する
2. 上記 FIR ローパスフィルタで生じた信号の群遅延（波形全体の時間遅れ）を補償する
3. ダウンサンプリング後の離散時間の各信号値を線形内挿（線形補間）により求める

また、心拍の推定範囲は 35–175 bpm ($f_{\min} = 0.58$ Hz 及び $f_{\max} = 3.5$ Hz) に制限した。NCF, CEP, 及び YIN では、1つの心拍値を算出する短時間信号長を $L = 1024$ 点 (51.2 s) とし、これを $S = 256$ 点 (12.8 s) ずつシフトした。また、MUS では、計算時間の都合から、 $L = 1024$ 点及び $S = 1024$ 点と設定した。従って、NCF, CEP, 及び YIN の 3 手法では推定心拍値が 12.8 s 毎に 1 つ（周波数で表すと 78.125 mHz）、MUS では 51.2 s 毎に 1 つ（周波数で表すと 19.53125 mHz）得られるような心拍推定となっている。さらに、YIN で用いる閾値は $\alpha = 0.1$ とし、MUS で用いる目的信号中の正弦波の個数は $N = 1$ とした。なお、いずれの手法に対しても、心拍値を推定した後にその時間変化をスムージングするため、長さ 12 点の移動中央値フィルタを適用した。

4.2 節で述べたとおり、この ECG から RRI を算出して得られる心拍の正解値の時間変化を Figs. 4.1 及び 4.2 にそれぞれ示す。但し、Fig. 4.1 は、心拍値の周波数を NCF, CEP, 及び

YIN の推定値に合わせるために、正解値を周波数 78.125 mHz にダウンサンプルしたグラフであり、Fig. 4.2 は MUS の推定値に合わせるために、正解値を周波数 19.53125 mHz にダウンサンプルしたグラフである。従って、PPG に対して基本周波数推定を適用した結果が、この Figs. 4.1 及び 4.2 にどの程度近いかが評価対象となる。但し、データ 1 は 2 章で述べたように時間同期が取られていないため、評価結果は参考に留める。

ここで、Table 4.1 は、データ 1 に対し各手法を適用した際の評価値を、それぞれ評価指標別にまとめたものである。いずれの手法も誤差が小さいほど精度が高いことを示す。なお、4 手法で最も誤差が小さい値（精度の高い値）を太字で示している。

4.3.2 実験結果比較

24 時間の PPG に各手法を適用して得られる心拍推定結果を Fig. 4.3 に示す。いずれの手法も睡眠時（横軸 1:00～7:00 及び 12:30-14:00 付近）は尤もらしい心拍が高精度で推定できていることが確認できるが、通常活動時は値を確認することが難しく、推定が困難であった。しかしながら、Fig. 4.3 (a) の NCF に着目すると、通常活動時のスパイクが多く生じてはいるものの、スパイクの中に 50–60 bpm 付近の信号値が複数確認できる。従って、再度移動中央値フィルタやローパスフィルタを適用すれば通常活動時もある程度推定ができるようになる可能性があると考えられる。また、Fig. 4.3 (c) の YIN では睡眠時のスパイクが最も少なく、50–60 bpm といった尤もらしい心拍が高精度で推定できていることが確認できる。Table 4.1 からは、CEP がすべての評価指標において、最も誤差が小さいことが確認できる。

4.4 データ 2 に対する実験結果

4.4.1 実験条件

心拍推定において、PPG のサンプリング周波数 f_s を測定時の 512 Hz から 20 Hz にダウンサンプリングして解析した。また、データ 1 の場合と同様に、ECG から RRI を算出して得られる心拍の正解値の時間変化を Figs. 4.4 及び 4.5 に示す。その他の実験条件は 4.3.1 節と同様である。

4.4.2 実験結果比較

24 時間の PPG の手法毎の心拍推定結果を Fig. 4.6 に示す。データ 1 の場合と同様に、いずれの手法も睡眠時（横軸 1:00～7:00 付近）は尤もらしい心拍が高精度で推定できていることが確認できるが、通常活動時にはデータ 1 と比べて、心拍推定精度が低下しているこ

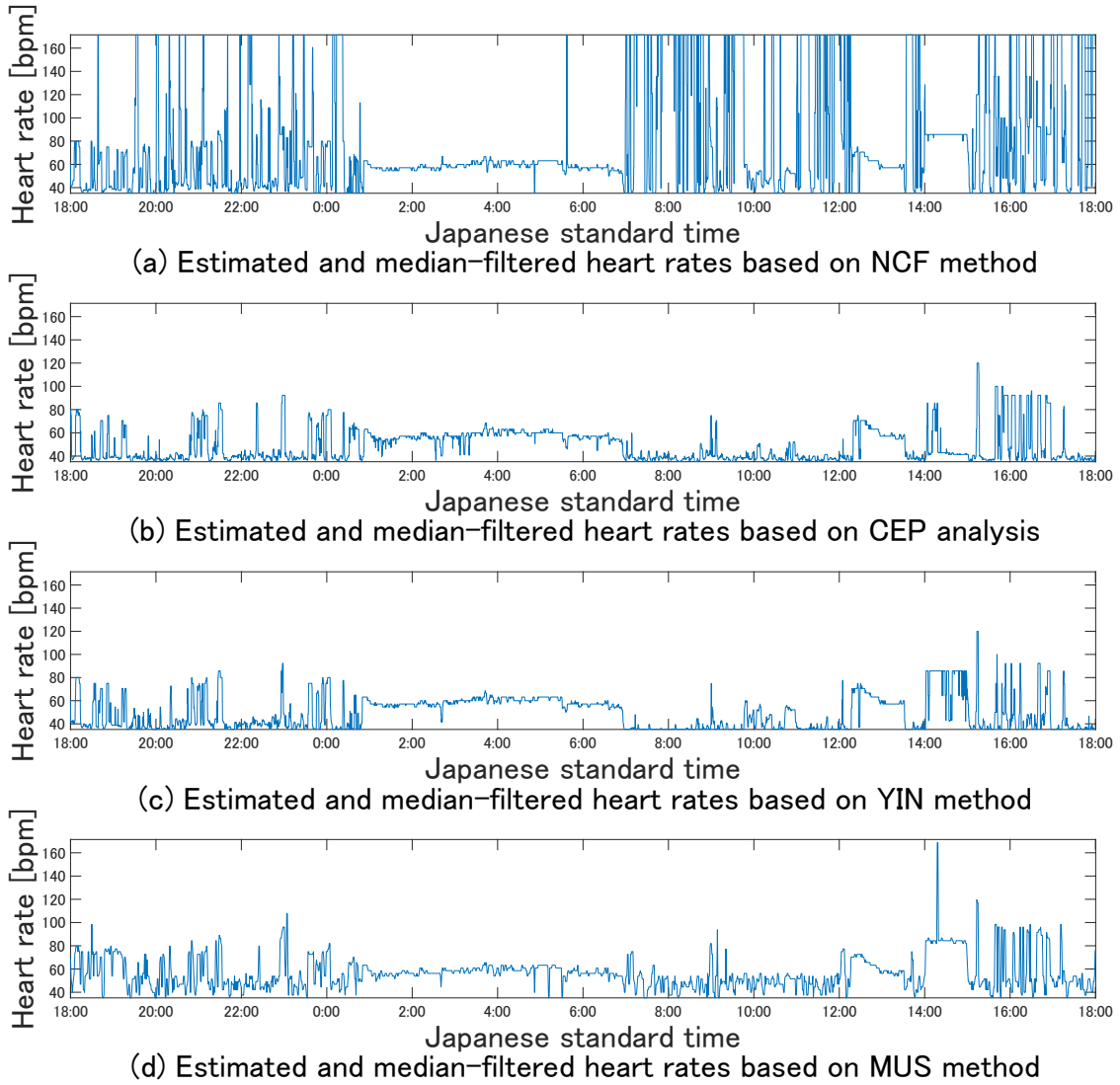


Fig. 4.3. Estimated heart rates using (a) NCF, (b) CEP, (c) YIN, and (d) MUS methods for Data 1.

とが確認できる。そのため、睡眠時の心拍における各手法の特徴がデータ 1 の場合と同様に確認できるが、運動時の心拍は全ての手法において推定下限として設定している 35 bpm となっている。Table 4.2 からは、YIN がすべての評価指標において、最も誤差が小さいことが確認できる。

4.5 データ 3 に対する実験結果

4.5.1 実験条件

心拍推定において、PPG のサンプリング周波数 f_s を測定時の 1024 Hz から 20 Hz にダウンサンプリングして解析した。データ 1 の場合と同様に、ECG から RRI を算出して得られる

Table 4.1. Evaluation scores of each method for Data 1

Evaluation method	NCF	CEP	YIN	MUS
MSE	3156.9242	345.3259	364.6778	399.1924
RMS	56.1865	18.5829	19.0965	19.9798
GPE ($p = 0.10$) [%]	74.2476	64.7798	67.4026	67.1209
GPE ($p = 0.15$) [%]	62.9207	48.7183	51.7854	52.9621
GPE ($p = 0.20$) [%]	55.0482	35.6794	39.9318	43.3649
GPE ($p = 0.25$) [%]	49.5923	26.1668	30.9676	37.0853
GPE ($p = 0.50$) [%]	36.4272	10.5645	11.7203	19.609
FPE ($p = 0.10$)	3.1821	2.7819	2.9654	3.0622
FPE ($p = 0.15$)	4.4632	4.0477	4.1604	4.3734
FPE ($p = 0.20$)	5.4679	5.0816	5.1237	5.3771
FPE ($p = 0.25$)	6.3191	5.9832	5.9548	6.1445
FPE ($p = 0.50$)	9.6381	8.9209	9.3606	9.2844

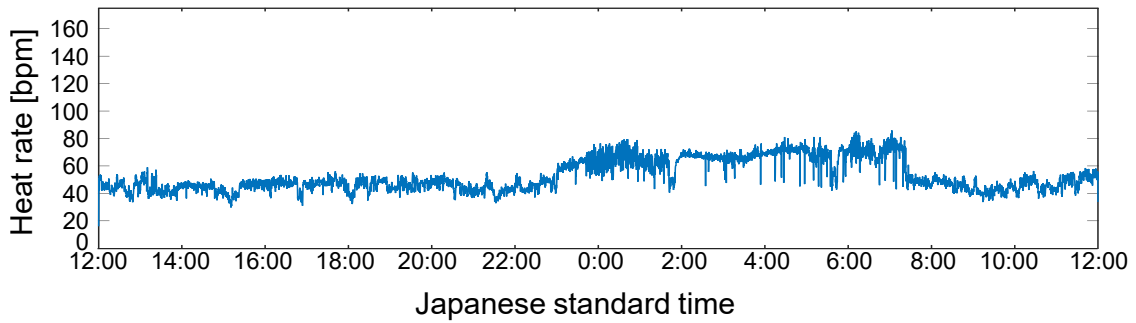


Fig. 4.4. Calculated heart rate obtained from ECG signal (Data 2) resampled at 78.125 mHz and used as reference.

心拍の正解値の時間変化を Figs. 4.7 及び 4.8 に示す．その他の実験条件は 4.3.1 節と同様である．

4.5.2 実験結果比較

24 時間の PPG の手法毎の心拍推定結果を Fig. 4.9 に示す．データ 1 及びデータ 2 の場合と同様に，23:50-06:10 の睡眠時の心拍はある程度の精度で推定できていることが分かるが，データ 2 と同様に運動時の心拍推定においてはデータ 1 と比べて，心拍推定精度が低下している．これは，クリッピング対策として用いた遮光テープの取り付けによって有用な信号が減衰したことで，推定に必要とする情報が減少したと考えられる．Table 4.3 からは，YIN がほぼすべての評価指標において，最も誤差が小さいことが確認できる．特に，FPE において，デー

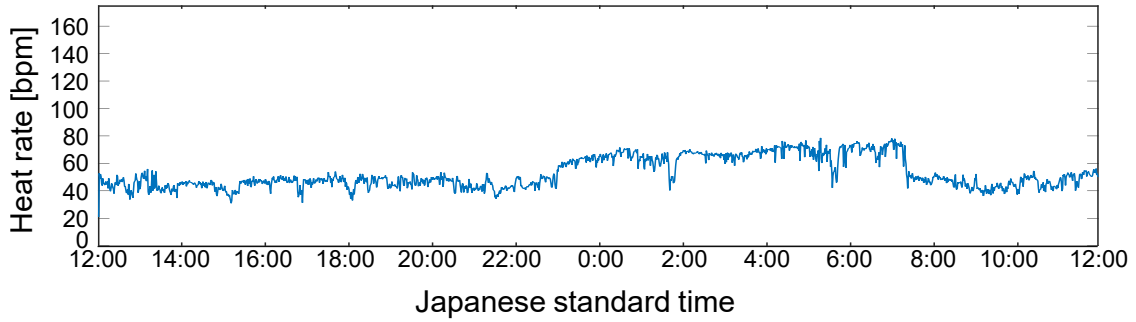


Fig. 4.5. Calculated heart rate obtained from ECG signal (Data 2) resampled at 19.53125 mHz and used as reference.

Table 4.2. Evaluation scores of each method for Data 2

Evaluation method	NCF	CEP	YIN	MUS
MSE	3670.8552	202.0897	191.2354	174.7788
RMS	60.5876	14.2158	13.8288	13.2204
GPE ($p = 0.10$) [%]	83.0411	80.3858	82.8155	66.1955
GPE ($p = 0.15$) [%]	72.9925	64.9242	69.5479	49.7644
GPE ($p = 0.20$) [%]	61.1168	46.6205	52.1425	37.3969
GPE ($p = 0.25$) [%]	50.7735	29.5244	34.7813	25.2061
GPE ($p = 0.50$) [%]	31.6045	2.1499	1.9437	4.1814
FPE ($p = 0.10$)	2.8571	2.3870	2.5519	2.8096
FPE ($p = 0.15$)	4.1247	3.3468	3.5083	4.3189
FPE ($p = 0.20$)	5.4023	4.3289	4.4001	5.6323
FPE ($p = 0.25$)	6.4945	5.0739	5.0041	7.0519
FPE ($p = 0.50$)	10.6511	8.8617	8.0574	10.5721

タ 1 及びデータ 2 と比べて YIN 法の誤差が小さいと言える。

4.6 本章のまとめ

本章では、本研究で行った実験内容とその結果について説明した。データ 2 及びデータ 3 と比較して、データ 1 では睡眠時の心拍が最も高精度に推定できた。これは、実験 2 及び 3 において、測定電圧がダイナミックレンジ外となりクリッピングすることへの対策のため、センサに装着した遮光テープの影響により、推定に必要な信号の成分が減衰していることが想定される。さらに、データ 1 では、CEP が最も誤差が小さいのに対して、データ 2 及び 3 では、YIN が最も誤差が小さいことが確認された。データ 2 及びデータ 3 では、運動時の心拍がどの手法においても推定できていない。そのため、睡眠時のみの心拍推定誤差の比較となっており、こ

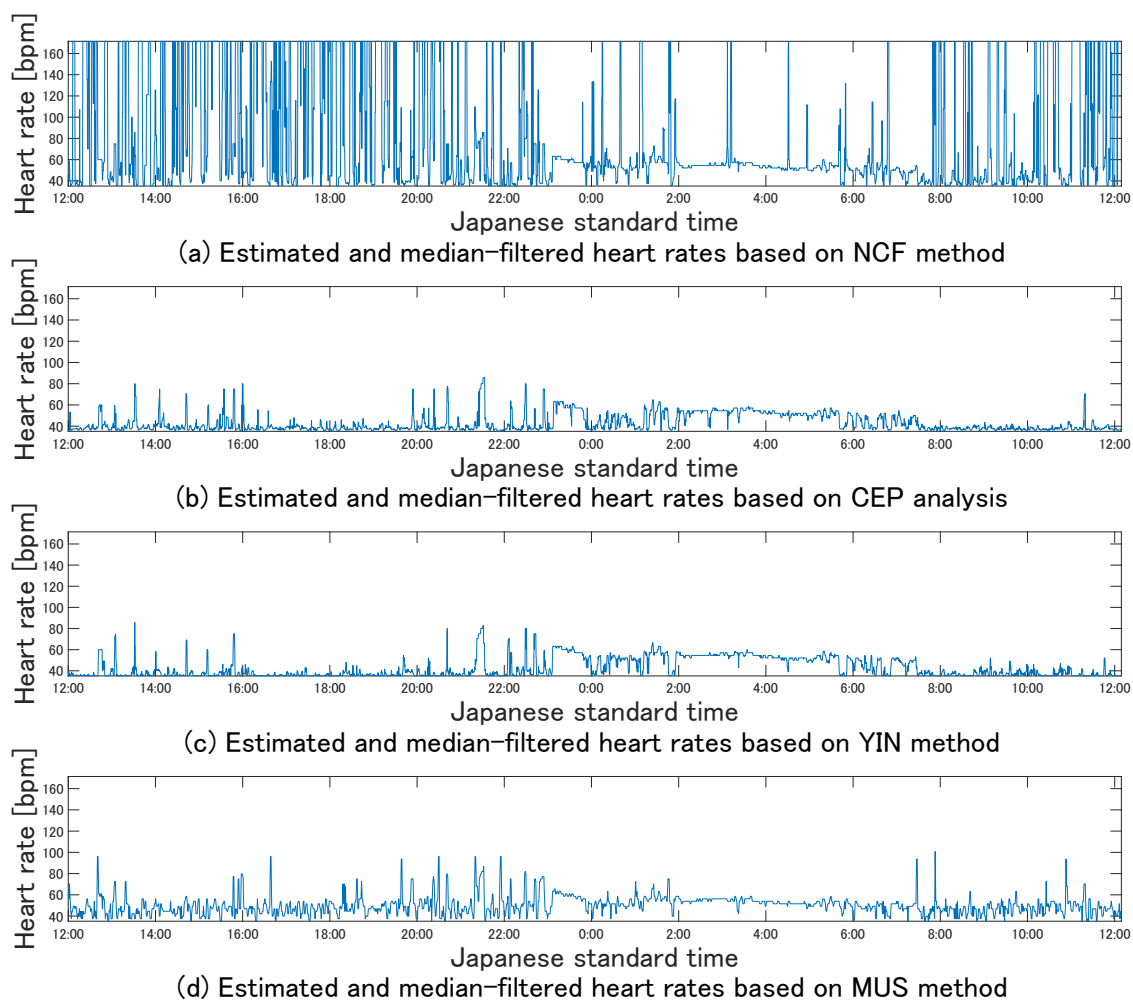


Fig. 4.6. Estimated heart rates using (a) NCF, (b) CEP, (c) YIN, and (d) MUS methods for Data 2.

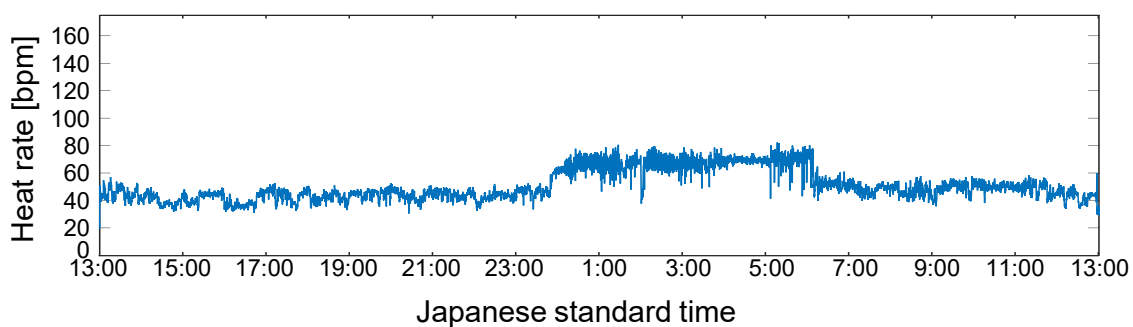


Fig. 4.7. Calculated heart rate obtained from ECG signal (Data 3) resampled at 78.125 mHzs and used as reference.

のような評価は FPE に強く反映される。運動時も加味した場合は、CEP が最も心拍推定精度が高く、睡眠時のみの場合は、YIN が最も心拍推定精度が高いと言える。このように、これら

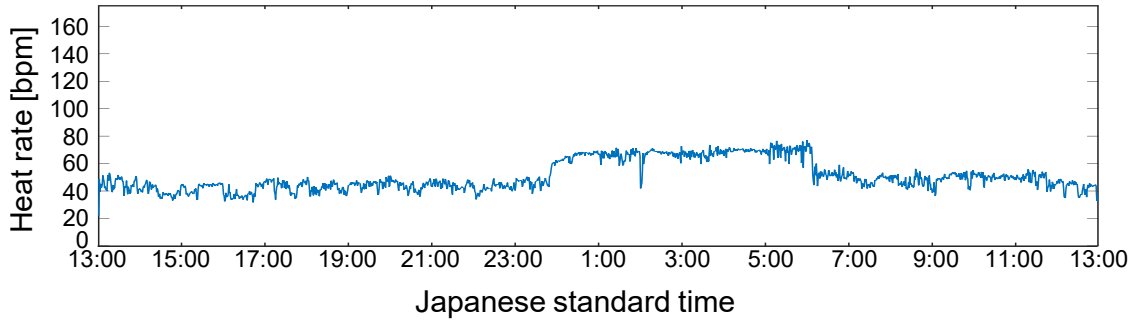


Fig. 4.8. Calculated heart rate obtained from ECG signal (Data 3) resampled at 19.53125 mHz and used as reference.

Table 4.3. Evaluation scores of each method for Data 3

Evaluation method	NCF	CEP	YIN	MUS
MSE	3522.7394	189.7079	172.9414	187.3802
RMS	59.3526	13.7734	13.1507	13.6887
GPE ($p = 0.10$) [%]	81.8384	76.0503	78.2101	70.4734
GPE ($p = 0.15$) [%]	72.7057	61.5385	65.5769	56.3314
GPE ($p = 0.20$) [%]	62.8478	44.9704	49.6893	42.7219
GPE ($p = 0.25$) [%]	51.1545	29.7485	31.4941	27.6331
GPE ($p = 0.50$) [%]	33.3629	2.4704	1.8935	5.4438
FPE ($p = 0.10$)	2.5807	2.3321	2.1327	2.6832
FPE ($p = 0.15$)	3.8229	3.2602	2.9994	4.1109
FPE ($p = 0.20$)	5.1755	4.2039	3.9683	5.6988
FPE ($p = 0.25$)	6.8346	5.2454	5.2055	7.5571
FPE ($p = 0.50$)	10.4550	8.8731	8.0543	10.8009

の手法はいずれも一長一短があるため、活動シーンに合わせて手法をそれぞれ切り替えることができるようなシステムが開発されれば、より高精度な心拍推定が実現できると予想される。

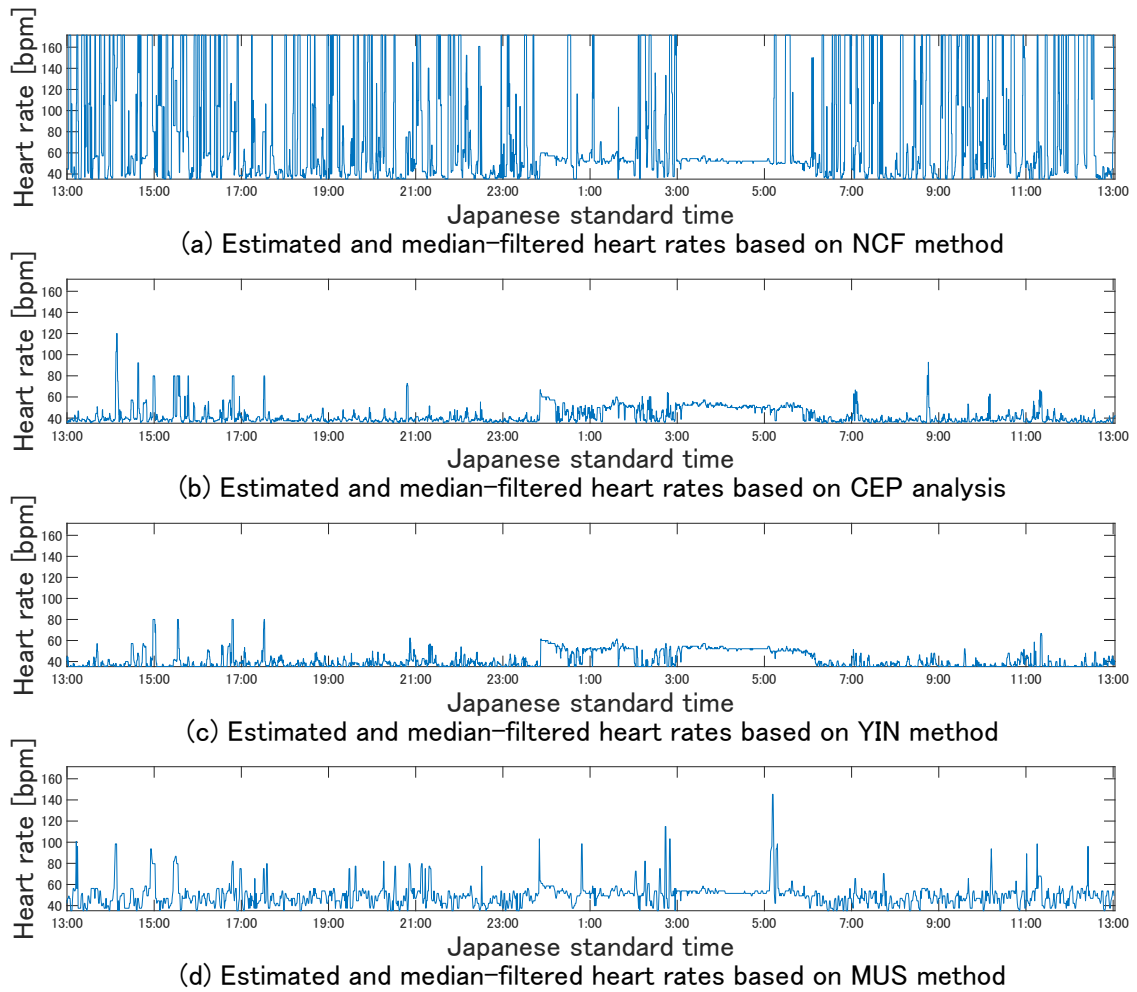


Fig. 4.9. Estimated heart rates using (a) NCF, (b) CEP, (c) YIN, and (d) MUS methods for Data 3.

第 5 章

結言

本論文では、付け爪型センサから得られる PPG に音響信号処理で代表的な基本周波数推定法である NCF 法、ケプストラム法、YIN 法、MUSIC 法の 4 つを適用し、生体信号処理分野でどの程度有効性があるか議論した。

1 章では、近年の基本周波数推定法に関する研究を紹介し、本論文の目的について述べた。2 章では、本研究で用いた付け爪型センサの概要について述べた。付け爪型センサを利用することで、これまでのウェアラブルセンサにおいて問題視されていた不快な装着感を改善でき、多点同時計測の実現や、ジェルネイルを用いることで長期間にわたる強固な固定も可能となる。3 章では、本研究で用いた 4 つの基本周波数推定法について述べた。いずれの手法も音楽信号処理の分野では基本周波数推定法として一般的に用いられる。4 章では、本研究で実施した実験の詳細な内容や結果について述べた。運動時も加味した場合はケプストラム法が、睡眠時のみ場合は YIN 法が最も精度が高く、手法毎に一長一短があることが分かった。

最後に今後の課題を述べる。現在の 4 手法では、生体信号処理分野への応用はある程度可能ではあるが、解析における推定精度が低いため、実用的であるとは言い難い。そこで、活動シーンに合わせて手法をそれぞれ切り替えることができるようなシステムを開発し、より精度を高めていく必要があると考えられる。

謝辞

本論文は、香川高等専門学校電気情報工学科北村研究室にて行われた研究に基づくものです。本研究を進めるにあたり、多くの方々にご指導ご鞭撻を賜りました。まず、日頃から研究の進み具合を気にかけていただき、不慣れな言語に途方に暮れる私に、豊富な知識と経験の下、熱心なご指導と適切な助言をくださった北村大地講師に心より感謝申し上げます。ありがとうございました。また、共同研究者である、機械電子工学科の石井耕平准教授には、生体信号処理分野や付け爪型センサに関する知識のご教授並びに、研究に必要なデータの提供など、不慣れな分野のサポートを賜り無事に修了することができそうです。ここに感謝いたします。そして、研究室のメンバーには常に刺激的な議論を頂き、時に研究の助言や激励をいただきながら精神的にも支えられました。ありがとうございます。最後に、本研究ならびに学業全般にわたって経済的・心身的に支援して下さる家族に深く感謝し、御礼申し上げます。

参考文献

- [1] K. Ishii and N. Hiraoka, “Nail tip sensor: toward reliable daylong monitoring of heart rate,” *Trans. Electrical and Electronic Eng.*, vol. 15, no. 6, pp. 902–908, 2020.
- [2] B. S. Atal, “Automatic speaker recognition based on pitch contours,” *J. Acoust. Soc. of Am.*, vol. 52, no. 6B, pp. 1687–1697, 1972.
- [3] IEEE Acoust., Speech, and Signal Process. Society, Digital Signal Process. Committee, *Programs for Digital Signal Processing*, IEEE Press, New York, 1979.
- [4] A. Cheveigne and H. Kawahara, “YIN, a fundamental frequency estimator for speech and music,” *J. Acoust. Soc. of Am.*, vol. 111, no. 4, pp. 1917–1930, 2002.
- [5] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, 1986.
- [6] Apple watch, <https://support.apple.com/ja-jp/guide/watch/apda88aefe4c/watchos>
- [7] ヘルスパッチ MD, <https://www.omron.com/jp/ja/news/2015/11/h1105.html>
- [8] M. G. Christensen, S. H. Jensen, S. V. Andersen, and A. Jakobsson, “Subspace-based fundamental frequency estimation,” *Proc. EUSIPCO*, pp. 637–640, 2004.
- [9] M. G. Christensen, A. Jakobsson, and S. H. Jensen, “Joint high-resolution fundamental frequency and order estimation,” *IEEE Trans. Audio, Speech, and Lang. Process.*, vol. 15, no. 5, pp. 1635–1644, 2007.
- [10] L. R. Rabiner, M. Cheng, A. Rosenberg, and C. McGonegal, “A comparative performance study of several pitch detection algorithms,” *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 24, no. 5, pp. 399–418, 1976.
- [11] 森勢将雅, “基本は検出に基づく F0 推定法の対雑音性向上,” *情報処理学会研究報告 IPSJ SIG Technical Report*, vol. 2016–SLP–110, no. 5, 2016.